

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 13:45:01 ; Search time 400 Seconds
(without alignments)
354.336 Million cell updates/sec

Title: US-10-073-064-2

Perfect score: 27
Sequence: 1 GGTCACCTACCGCGGTGAACGGGCATT 27

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4134886 segs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	27	100.0	27	2	AAQ42309	Aa42309 Sequence
2	27	100.0	27	2	AAx8057	Aax8057 Vitamin D
3	25	92.6	57	4	AAx8560	Aax8560 Human ost
4	23	85.2	53	2	AAV18523	Aav18523 Primer VD
5	23	85.2	53	2	AAx16604	Aax16604 Human vit
6	23	85.2	57	4	AAx8560	Aax8560 Human ost
7	23	85.2	57	4	AAx8563	Aax8563 Human ost
8	23	85.2	57	4	AAx8563	Aax8563 Human ost
9	23	85.2	301	4	AAx8563	Aax8563 Human ost
10	23	85.2	301	4	AAx8563	Aax8563 Human ost
11	23	85.2	565	8	AAx8563	Aax8563 Human ost
12	23	85.2	565	8	AAx8563	Aax8563 Human ost
13	23	85.2	597	5	AAx8563	Aax8563 Human ost
14	23	85.2	597	5	AAx8563	Aax8563 Human ost
15	23	85.2	2650	12	AAx8563	Aax8563 Human ost
16	23	85.2	2657	8	AAx8563	Aax8563 Human ost
17	23	85.2	2657	8	AAx8563	Aax8563 Human ost
18	23	85.2	2657	8	AAx8563	Aax8563 Human ost
19	23	85.2	2657	8	AAx8563	Aax8563 Human ost
20	23	85.2	2657	8	AAx8563	Aax8563 Human ost
21	23	85.2	2657	8	AAx8563	Aax8563 Human ost

C	22	19.6	72.6	13203	4	ABJ30124	Abj30124 Drosophil
C	23	19	70.4	6846	4	ABJ18560	Abj18560 Drosophil
C	24	19	70.4	12631	4	ABJ17964	Abj17964 Drosophil
C	25	18.2	67.4	132	12	ACH86244	Ach86244 Human gen
C	26	18.2	67.4	545	12	ACH72488	Ach72488 Human gen
C	27	18.2	67.4	721	4	AA159837	Aa159837 Human pol
C	28	18.2	67.4	995	6	AAJ26472	Aad26472 Human kin
C	29	18.2	67.4	1223	4	AAJ27222	Aad27222 Human enco
C	30	18.2	67.4	1223	10	ADB93400	Adv93400 Human gen
C	31	18.2	67.4	1440	10	ABV75301	Abv75301 NOVEL hum
C	32	18.2	67.4	1452	4	AAJ14324	Aad14324 Human 278
C	33	18.2	67.4	1452	10	AAJ61170	Aad61170 Human 278
C	34	18.2	67.4	1579	10	AAJ53813	Aad53813 Human cod
C	35	18.2	67.4	1581	8	AAJ53173	Aad53173 Human kin
C	36	18.2	67.4	1619	10	ADB63503	Adb63503 Human GDN
C	37	18.2	67.4	1623	10	ADCI10109	Adci10109 Human NOV
C	38	18.2	67.4	1778	4	AA158051	Aa158051 Human pol
C	39	18.2	67.4	1967	10	ADCI10107	Adci10107 Human NOV
C	40	18.2	67.4	5340	6	ABJ32234	Abj32234 Human imm
C	41	18	66.7	18	8	ABZ58132	Abz58132 Human ost
C	42	18	66.7	624	11	ACH96158	Ach96158 Klebsiell
C	43	18	66.7	724	4	AAH08666	Aah08666 Human GDN
C	44	18	66.7	768	11	ACH96135	Ach96135 Klebsiell
C	45	18	66.7	1063	8	ACA35059	Aca35059 Prokaryot

ALIGNMENTS

RESULT 1
ID AAQ42309 standard; cDNA; 27 BP.

XX AAQ42309;
XX 25-MAR-2003 (revised)
DT 09-JAN-2003 (revised)
DT 26-AUG-1993 (first entry)

DE Sequence associated with myogenic vector systems.

XX Myogenic vector system; expression; myogenic tissue; vector; vaccine;
KM insulin-like growth factor; IGF-II; insulin growth factor; milk; meat;
KM binding protein; growth hormone; growth hormone releasing hormone;
KM apolipoprotein; A-I; glycogen phosphorylase; alpha-1-antitrypsin; human;
KM animal; tissue culture; gene replacement; dytrophin; IGF-I;
KM muscle atrophy; spinal column; injury; neuromuscular disease; MSV;
KM vitamin D; regulatory element; VDRE; herpes simplex virus; HSV; ss.

XX Unidentified.
OS
XX
XX WO9309236-A1.
XX
XX 13-MAY-1993.
XX
XX 03-NOV-1992; 92WO-US009353.
XX
XX 06-NOV-1991; 91US-00789919.
XX
XX (BAYU) BAYLOR COLLEGE MEDICINE.
XX
XX Schwartz RJ, Demayo FJ, Omalley BW;
XX
XX WPI; 1993-167701/20.
XX
XX Myogenic vector systems for expressing nucleic acid sequences - useful in
PT gene replacement therapy to treat muscle atrophy, atherosclerotic
PT cardiovascular conditions, etc.
XX
XX Disclosure; Page 29-30; 56pp; English.
XX
XX This sequence represents a target sequence which was designed to contain
XX synthesized multimers of the vitamin D regulatory element (VDRE). This

CC target was linked to the herpes simplex virus (HSV) promoter and used in
 CC a myogenic vector system (MSV) of the invention. The MSV are capable of
 CC expressing any specific nucleic acid sequence in myogenic tissue. This
 CC involves a regulatable myogenic vector system. The vectors can be used
 CC for the expression of eg. insulin-like growth factor (IGF-I, IGF-II,
 CC insulin growth factor binding protein, growth hormone, growth hormone
 CC releasing hormone, apolipoprotein A-I, glycogen phosphorylase, alpha-1-
 CC antitrypsin or dystrophin. They can be used for the supply of polypeptide
 CC into human, animal or tissue culture, for gene replacement or for vaccine
 CC production. They can be used for eg. treating muscle atrophy in aging
 CC humans or induced by spinal column injuries or neuromuscular diseases,
 CC preventing or treating growth disease. When the vector encode growth
 CC hormone they can be used for increasing milk or meat production in
 CC animals. (Updated on 09-JAN-2003 to add missing OS field.) (Updated on 25
 CC -MAR-2003 to correct PN field.)

CC Sequence 27 BP; 5 A; 6 C; 11 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 27; DB 2; Length 27;

Best Local Similarity 100.0%; Pred. No. 0.018;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGCGGTGAACGGGGCATT 27

DB 1 GGTGACTCACCAGCGGTGAACGGGGCATT 27

RESULT:2
 AAX88057
 ID AAX88057 standard; cDNA; 27 BP.

AC AAX88057;

DT 08-SEP-1999 (first entry)

DE Vitamin D regulatory element (VDRE) cDNA.

KM Plasmid pIG0100A; plasmid pIG035; expression vector; treatment; disease;

KM RNA stability element; gene therapy; muscle atrophy; neurological;

KM muscular disease; systemic disease; aging; trophic factor; haemophilia;

KM clotting factor; atherosclerosis; atherosclerotic; cardiovascular;

KM cerebrovascular; peripheral-vascular disease; hormone deficiency;

KM diabetes; transgenic animal; carcinogen; regulatory element;

KM livestock improvement; immune response; ss.

OS Unidentified.

XX US5925564-A.

XX 20-JUL-1999.

XX 07-JUN-1995; 95US-00472809.

XX 06-NOV-1991; 91US-00789919.

XX 09-MAR-1994; 94US-00209846.

XX (BAYU) BAYLOR COLLEGE MEDICINE.

XX O'malley BM, Demayo FJ, Schwartz RJ;

XX WPI, 1999-418276/35.

XX New expression vector system useful for gene therapy.

XX Disclosure; Col 39-40; 67pp; English.

CC neurological, muscular or systemic disease, aging by causing tissues to
 CC express trophic factors, haemophilia by causing tissues to express and
 CC secrete clotting factor into the circulation, atherosclerosis and
 CC atherosclerotic cardiovascular, cerebrovascular or peripheral-vascular
 CC disease by causing tissues to express factors involved in tissue
 CC metabolism. They can be used to replace genes of inherited genetic
 CC defects or acquired hormone deficiencies e.g. diabetes. To transform
 CC cells to produce particular proteins or RNA in vitro. To create
 CC transgenic animals which can be used for research into human diseases,
 CC assessing novel therapeutic methods, assessing the effect of chemical and
 CC physical carcinogens and for studying the effect of genes and genetic
 CC regulatory elements or livestock improvement. They can be used to induce
 CC an immune response. These vectors provide controlled expression of the
 CC genes they carry and produce a significantly high level of expression.
 CC Using 3'UTR sequences reduces the decay rates of the mRNAs encoded by the
 CC vectors which causes increased expression

CC Sequence 27 BP; 5 A; 6 C; 11 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 27; DB 2; Length 27;

Best Local Similarity 100.0%; Pred. No. 0.018;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGCGGTGAACGGGGCATT 27

DB 1 GGTGACTCACCAGCGGTGAACGGGGCATT 27

RESULT 3
 AAF85650
 ID AAF85650 standard; DNA; 57 BP.

AC AAF85650;

DT 25-JUN-2001 (first entry)

DE Human osteocalcin vitamin D response element #2.

KM Marmoser; vitamin D response element binding protein; VDRE-BP; lymphoma;

KM vitamin D resistance; osteoporosis; hypercalcaemia; vitamin D toxicity;

KM glucocorticoid mediated disorder; granuloma forming disease;

KM vitamin D intoxication; steroid hormone hypersecretion; gene therapy; ds.

XX Homo sapiens.

OS

XX Key Location/Qualifiers

XX misc_feature 14..19

XX /*tag= a

XX /note= "response element half site"

XX misc_feature 23..29

XX /*tag= b

XX /note= "response element half site"

XX WO200121649-A2.

XX 29-MAR-2001.

XX 20-SEP-2000; 2000WO-US025844.

XX 22-SEP-1993; 99US-00400967.

XX (CEDA-) CEDARS SINAI MEDICAL CENT.

XX Adams JS;

XX WPI, 2001-308082/32.

XX New vitamin D response element-binding protein (VDRE-BP) useful in
 XX modifying vitamin D receptor activity, in producing anti-VDRE-BP
 XX antibodies, in identifying agonists and antagonists of the protein, or in
 XX gene therapy.

XX Example; Page 46; 81pp; English.

XX The present invention provides the protein and coding sequences for two
 CC vitamin D response element binding proteins (VDRE-BPs) from the marmoset.
 CC These confer vitamin D resistance on the primate, a characteristic which
 CC is associated with high circulating levels of other steroid hormones. The
 CC sequences provided by the invention can be used to identify treatments
 CC for osteoporosis, hypercalcaemia, vitamin D intoxication, steroid hormone
 CC hypersecretion, glucocorticoid mediated disorders, lymphoma and granuloma
 CC forming diseases. The present sequence is an example of a vitamin D
 CC response element. Note: The present sequence is stated in the
 CC specification as being the same as that shown as SEQ ID NO: 6 on page 46
 CC (see AAF85639). However, the sequences differ at the last position
 CC XX
 SQ Sequence 57 BP; 12 A; 18 C; 17 G; 10 T; 0 U; 0 Other;

Query Match 92.6%; Score 25; DB 4; Length 57;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTCACCTACCGGGTGAACGGGGCA 25
 |||||
 DB 3 GGTCACCTACCGGGTGAACGGGGCA 27

RESULT 4
 AAV18523
 ID AAV18523 standard; DNA; 53 BP.
 XX
 AC AAV18523;

DT 11-JUN-1998 (first entry)
 XX
 DE Primer VDRE(OC) for rat vitamin D receptor cDNA.

KM Rat; vitamin D receptor; isoform protein; VDRI; VDR0; diagnosis;
 KW dominant negative receptor; signal transduction channel;
 KM bone density disorder; screening; PCR primer; ss.

OS Synthetic.
 XX Rattus rattus.
 XX
 PN M09747172-A1.

PD 18-DEC-1997.

PF 10-JUN-1997; 97WO-1B000947.

PR 10-JUN-1996; 96JP-00194179.

PA (CHUS) CHUGAI SEIYAKU KK.

PI Kato S, Ueno K;

DR WPI; 1998-051917/05.

PT DNA encoding a vitamin D receptor isoform protein - useful for bone
 PT density determination and for screening substances for vitamin D
 PT activity.

PS Example 3; Page 18; 46pp; Japanese.

CC The present sequence was used in the preparation of the cDNA encoding the
 CC rat vitamin D receptor isoform protein (VDRI). The isoform differs from
 CC the normal receptor (VDR0) in having the vitamin D response element
 CC curtailed by 86 residues, and having an extra 19 residues inserted at the
 CC C-terminal of this element. It acts as a dominant negative receptor in
 CC the vitamin D signal transduction channel. The isoform protein can be
 CC used to diagnose bone density disorders, and screen for substances having
 CC potential vitamin D-like activity

SQ Sequence 53 BP; 8 A; 13 C; 21 G; 11 T; 0 U; 0 Other;

Query Match 85.2%; Score 23; DB 2; Length 53;

Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTCACCTACCGGGTGAACGGGG 23
 |||||
 DB 17 GGTCACCTACCGGGTGAACGGGG 39

RESULT 5
 AAX16604
 ID AAX16604 standard; DNA; 53 BP.
 XX
 AC AAX16604;

DT 25-APR-1999 (first entry)

DE Human vitamin D receptor isoform protein VDRI PCR primer OC.

KM Vitamin D receptor; VDR; bone density; retinoic acid derivative; steroid;
 KW bone calcium regulator; immunosuppressant; anticancer; PCR primer; ss.

OS Synthetic.
 XX Homo sapiens.

PN M09856908-A1.

PD 17-DEC-1998.

PF 13-JUN-1997; 97WO-JP002052.

PR 13-JUN-1997; 97WO-JP002052.

PA (CHUS) CHUGAI SEIYAKU KK.

PI Kato S, Ueno K;

DR WPI; 1999-080898/07.

PT Gene for vitamin D receptor isoform protein which blocks vitamin D signal
 PT pathway - and production of recombinant isoform protein using it, for
 PT bone density assay and for screening compounds for vitamin D activity.

PS Example 3; Page 18; 47pp; Japanese.

CC The present sequence represents a PCR primer for human vitamin D receptor
 CC (VDR) isoform protein VDRI. VDRI differs from the canonical form of VDR
 CC (VDR0) by lacking the amino acid sequence encoded by exon 9, but instead
 CC substitutes a short sequence (GTPGSESRIDGHTGDCS in the rat protein)
 CC encoded by the 5'-end of intron 8. VDRI has a dominant negative effect on
 CC the vitamin D signalling pathway. The isoform protein can be used for
 CC determining bone density, and for the screening of compounds (e.g.
 CC steroids and retinoic acid derivatives) for vitamin D activity (e.g. as
 CC bone calcium regulators, immunosuppressants or anticancer agents)

SQ Sequence 53 BP; 8 A; 13 C; 21 G; 11 T; 0 U; 0 Other;

Query Match 85.2%; Score 23; DB 2; Length 53;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTCACCTACCGGGTGAACGGGG 23
 |||||
 DB 17 GGTCACCTACCGGGTGAACGGGG 39

RESULT 6
 AAF85650/C
 ID AAF85650 standard; DNA; 57 BP.
 XX
 AC AAF85650;

DT 25-JUN-2001 (first entry)

XX

```

DE Human osteocalcin vitamin D response element #2.
XX
KM Marmoset; vitamin D response element binding protein; VDRE-BP; lymphoma;
KM vitamin D resistance; osteoporosis; hypercalcaemia; vitamin D toxicity;
KM glucocorticoid mediated disorder; granuloma forming disease;
KM vitamin D intoxication; steroid hormone hypersecretion; gene therapy; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_feature 14..19
FT /tag= a
FT /note= "response element half site"
FT misc_feature 23..29
FT /tag= b
FT /note= "response element half site"
XX
XX WO200121649-A2.
XX
XX 29-MAR-2001.
XX
XX 20-SEP-2000; 2000WO-US025844.
XX
XX 22-SEP-1999; 99US-00400967.
XX
XX (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
XX Adams JS;
XX
XX WPI; 2001-308082/32.
XX
XX New vitamin D response element-binding protein (VDRE-BP) useful in
XX modifying vitamin D receptor activity, in producing anti-VDRE-BP
XX antibodies, in identifying agonists and antagonists of the protein, or in
XX gene therapy.
XX
XX Example; Page 46; 81pp; English.
XX
XX The present invention provides the protein and coding sequences for two
XX vitamin D response element binding proteins (VDRE-BPs) from the marmoset.
XX These confer vitamin D resistance on the primate, a characteristic which
XX is associated with high circulating levels of other steroid hormones. The
XX sequences provided by the invention can be used to identify treatments
XX for osteoporosis, hypercalcaemia, vitamin D intoxication, steroid hormone
XX hypersecretion, glucocorticoid mediated disorders, lymphoma and granuloma
XX forming diseases. The present sequence is an example of a vitamin D
XX response element. Note: The present sequence is stated in the
XX specification as being the same as that shown as SEQ ID NO: 6 on page 46
XX (see AAF85639). However, the sequences differ at the last position
XX
XX Sequence 57 BP; 12 A; 18 C; 17 G; 10 T; 0 U; 0 Other;
XX
XX Query Match 85.2%; Score 23; DB 4; Length 57;
XX Best Local Similarity 100.0%; Pred. No. 1.2;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GGTGACTCACCAGGTGAACGGGG 23
XX |||||
XX Db 55 GGTGACTCACCAGGTGAACGGGG 33
XX
XX RESULT 7
XX ID AAF85639 standard; DNA; 57 BP.
XX
XX AAF85639;
XX
XX 25-JUN-2001 (first entry)
XX
XX Human osteocalcin vitamin D response element.
XX
XX Marmoset; vitamin D response element binding protein; VDRE-BP; lymphoma;
XX vitamin D resistance; osteoporosis; hypercalcaemia; vitamin D toxicity;
XX glucocorticoid mediated disorder; granuloma forming disease;
XX vitamin D intoxication; steroid hormone hypersecretion; gene therapy; ds.
XX

```

```

KM glucocorticoid mediated disorder; granuloma forming disease;
KM vitamin D intoxication; steroid hormone hypersecretion; gene therapy; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_feature 14..19
XX /tag= a
XX /note= "response element half site"
XX misc_feature 23..29
XX /tag= b
XX /note= "response element half site"
XX
XX WO200121649-A2.
XX
XX 29-MAR-2001.
XX
XX 20-SEP-2000; 2000WO-US025844.
XX
XX 22-SEP-1999; 99US-00400967.
XX
XX (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
XX Adams JS;
XX
XX WPI; 2001-308082/32.
XX
XX New vitamin D response element-binding protein (VDRE-BP) useful in
XX modifying vitamin D receptor activity, in producing anti-VDRE-BP
XX antibodies, in identifying agonists and antagonists of the protein, or in
XX gene therapy.
XX
XX Example; Page 76; 81pp; English.
XX
XX The present invention provides the protein and coding sequences for two
XX vitamin D response element binding proteins (VDRE-BPs) from the marmoset.
XX These confer vitamin D resistance on the primate, a characteristic which
XX is associated with high circulating levels of other steroid hormones. The
XX sequences provided by the invention can be used to identify treatments
XX for osteoporosis, hypercalcaemia, vitamin D intoxication, steroid hormone
XX hypersecretion, glucocorticoid mediated disorders, lymphoma and granuloma
XX forming diseases. The present sequence is an example of a vitamin D
XX response element. Note: The present sequence is stated in the
XX specification as being the same as that shown as SEQ ID NO: 6 in the
XX sequence listing (see AAF85630). However, the sequences differ at the
XX last position
XX
XX Sequence 57 BP; 11 A; 18 C; 18 G; 10 T; 0 U; 0 Other;
XX
XX Query Match 85.2%; Score 23; DB 4; Length 57;
XX Best Local Similarity 100.0%; Pred. No. 1.2;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GGTGACTCACCAGGTGAACGGGG 23
XX |||||
XX Db 3 GGTGACTCACCAGGTGAACGGGG 25
XX
XX RESULT 8
XX ID AAF85639/c
XX
XX AAF85639;
XX
XX 25-JUN-2001 (first entry)
XX
XX Human osteocalcin vitamin D response element.
XX
XX Marmoset; vitamin D response element binding protein; VDRE-BP; lymphoma;
XX vitamin D resistance; osteoporosis; hypercalcaemia; vitamin D toxicity;
XX glucocorticoid mediated disorder; granuloma forming disease;
XX vitamin D intoxication; steroid hormone hypersecretion; gene therapy; ds.
XX

```

```
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_feature 14..19
XX /tag= a
XX /note= "response element half site"
XX /tag= b
XX /note= "response element half site"
XX
XX PN WO00121649-A2.
XX PD 29-MAR-2001.
XX PF 20-SEP-2000; 2000WO-US025844.
XX PR 22-SEP-1999; 99US-00400967.
XX (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
XX PI Adams JS;
XX WPI; 2001-308082/32.
XX DR
XX PT New vitamin D response element-binding protein (VDRE-BP) useful in
XX modifying vitamin D receptor activity, in producing anti-VDRE-BP
XX antibodies, in identifying agonists and antagonists of the protein, or in
XX gene therapy.
XX
XX PS Example; Page 76; 81pp; English.
XX
XX CC The present invention provides the protein and coding sequences for two
XX vitamin D response element binding proteins (VDRE-BPs) from the marmoset.
XX These confer vitamin D resistance on the primate, a characteristic which
XX is associated with high circulating levels of other steroid hormones. The
XX sequences provided by the invention can be used to identify treatments
XX for osteoporosis, hypercalcaemia, vitamin D intoxication, steroid hormone
XX hypersecretion, glucocorticoid mediated disorders, lymphoma and granuloma
XX forming diseases. The present sequence is an example of a vitamin D
XX response element. Note: The present sequence is stated in the
XX specification as being the same as that shown as SEQ ID NO: 6 in the
XX sequence listing (see AAF85650). However, the sequences differ at the
XX last position
XX
XX SQ Sequence 57 BP; 11 A; 18 C; 18 G; 10 T; 0 U; 0 Other;
XX
XX Query Match 85.2%; Score 23; DB 4; Length 57;
XX Best Local Similarity 100.0%; Pred. No. 1.2;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GGTGACTCACCGGTGAACGGG 23
XX |||||
XX 56 GGTGACTCACCGGTGAACGGG 34
XX
XX RESULT 9
XX AAQ79120
XX ID AAQ79120 standard; DNA; 301 BP.
XX
XX AC AAQ79120;
XX
XX DT 25-MAR-2003 (revised)
XX DT 06-AUG-1995 (first entry)
XX
XX DE Osteocalcin gene 5' region.
XX
XX KM TNFR; tumor necrosis factor responsive element; osteocalcin; ss.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_RNA 229..237
XX /tag= a
```

```
FT /label= TNFR
FT /note= "Claim 4, page 42"
XX
XX PN WO9428124-A1.
XX
XX PD 08-DEC-1994.
XX
XX PF 18-MAY-1994; 94WO-US005659.
XX
XX PR 20-MAY-1993; 93US-00066372.
XX
XX PA (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
XX
XX PI Li Y, Stashenko P;
XX WPI; 1995-022796/03.
XX DR
XX PT New TNF-alpha responsive element and DNA-binding protein - used to
XX develop prods to inhibit or mediate TNF-alpha induced inhibition of gene
XX transcription.
XX
XX PS Disclosure; Page 32; 63pp; English.
XX
XX CC Tumor necrosis factor (TNF)-mediated down-regulation of the osteocalcin
XX gene is exerted through a TNF responsive element (TNFRE). Deletion
XX analysis localized the TNFRE to -523/-515 of the 5' flanking region of
XX the osteocalcin gene. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 301 BP; 41 A; 111 C; 84 G; 65 T; 0 U; 0 Other;
XX
XX Query Match 85.2%; Score 23; DB 2; Length 301;
XX Best Local Similarity 100.0%; Pred. No. 1.4;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GGTGACTCACCGGTGAACGGG 23
XX |||||
XX 240 GGTGACTCACCGGTGAACGGG 262
XX
XX RESULT 10
XX ABZ58131
XX ID ABZ58131 standard; DNA; 565 BP.
XX
XX AC ABZ58131;
XX
XX DT 22-Apr-2003 (first entry)
XX
XX DE Human osteocalcin promoter.
XX
XX KM Human; osteocalcin; promoter; bone; tumour; prostate cancer; metastasis;
XX gene therapy; diagnosis; prognosis; marker; cyostatic; ds.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_signal 74..91
XX /tag= a
XX /note= "vitamin D response element"
XX /tag= c
XX /note= "OSE2 element"
XX /tag= b
XX /note= "OSE1 element"
XX
XX PN WO003006621-A2.
XX
XX PD 23-JAN-2003.
XX
XX PF 12-JUL-2002; 2002WO-US022216.
XX
XX PR 13-JUL-2001; 2001US-0305360P.
XX
```

PA (UYVI-) UNIV VIRGINIA PATENT FOUND.
 XX
 XX Chung LWK, Yeung F;
 XX
 XX WPI; 2003-221733/21.
 XX
 XX Nucleic acid sequence for diagnosing, prognosing or treating calcified
 PT tumors and tissues, e.g. prostate cancer, comprises an improved
 PT recombinant human osteocalcin promoter activity.
 XX
 XX Claim 12; Page 51-52; 55pp; English.
 XX
 CC The present sequence is that of the human osteocalcin (hOC) promoter. An
 CC evaluation of the hOC promoter was conducted in which the functional
 CC hierarchy of the cis-acting elements OSB1, OSB2 and AP-1/VDRB (vitamin D
 CC response element) was defined in an androgen-independent human prostate
 CC cancer PC-3 cell line. By juxtaposing dimers of these 3 elements, a
 CC minimal hOC super-promoter (see AB258130) was produced. The super-
 CC promoter displays over 8-fold higher activity than the native full-length
 CC hOC promoter in a tissue-specific manner in PC-3 cells. Expression
 CC vectors comprising the hOC promoter sequences and transformed host cells
 CC are provided by the invention. In one embodiment, the hOC super-promoter
 CC is operably linked to a nucleic acid encoding a heterologous protein,
 CC ribozyme, dominant-negative or antisense RNA and used to deliver
 CC therapeutic genes to localized or disseminated tumors. The hOC promoter
 CC sequences can also be used to deliver therapeutic genes to fractured
 CC bones for bone repair. hOC promoter activation by extracellular matrices
 CC and soluble factors secreted by prostate cancer and bone cells in useful
 CC as a marker for the diagnosis and prognosis of prostate cancer
 XX
 SQ Sequence 565 BP; 101 A; 185 C; 170 G; 109 T; 0 U; 0 Other;
 Query Match 85.2%; Score 23; DB 8; Length 565;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTGACTCACC GGGTGAACGGGG 23
 DB 74 GGTGACTCACC GGGTGAACGGGG 96
 RESULT 11
 AAT90018
 ID AAT90018 standard; DNA; 597 BP.
 XX
 XX AAT90018;
 XX
 XX 20-NOV-1997 (first entry)
 XX
 DE Bases -575 to 22 of the human osteocalcin gene.
 XX
 XX Recombinant vector; human; osteocalcin; promoter; transcription;
 KW control region; determination; activity; vitamin D derivative;
 KW bone metabolism; regulation; ss.
 XX
 OS Homo sapiens.
 XX
 PN JP09182586-A.
 XX
 PD 15-JUL-1997.
 XX
 XX 29-DEC-1995; 95JP-00352817.
 XX
 XX 29-DEC-1995; 95JP-00352817.
 XX
 XX 29-DEC-1995; 95JP-00352817.
 XX
 PA (SRLS-) SRL KK.
 XX
 DR WPI; 1997-409886/38.
 XX
 PT Recombinant vector comprising vitamin D response element - and human
 PT osteocalcin gene promoter sequence and transcriptional control region,
 PT upstream of structural gene encoding detectable compound.
 XX

PS Claim 3; Page 5; 7pp; Japanese.
 XX
 XX A novel recombinant vector, comprises a vitamin D response element
 CC (VDRB), and a human osteocalcin gene promoter sequence and
 CC transcriptional control region, i.e. the present sequence, upstream of a
 CC structural gene encoding a detectable compound. The vector can be used to
 CC determine the activity of vitamin D derivatives as bone metabolism
 CC regulators, without using animals. A transformant containing the vector
 CC is incubated on a dish to confluence, and a vitamin D derivative added at
 CC a concentration of 10⁻¹¹ to 10⁻⁸ M. The cells are incubated for 2 days
 CC and recovered, and the yield of the detectable compound determined. When
 CC the vitamin D derivative is highly active the yield of the detectable
 CC compound is increased, as the osteocalcin gene promoter is activated
 XX
 SQ Sequence 597 BP; 107 A; 190 C; 187 G; 113 T; 0 U; 0 Other;
 Query Match 85.2%; Score 23; DB 2; Length 597;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTGACTCACC GGGTGAACGGGG 23
 DB 65 GGTGACTCACC GGGTGAACGGGG 87
 RESULT 12
 AAD14384
 ID AAD14384 standard; DNA; 597 BP.
 XX
 XX AAD14384;
 XX
 XX 01-NOV-2001 (first entry)
 XX
 DE Human osteocalcin promoter DNA.
 XX
 XX Human; antiinflammatory; recombinant human interleukin-4; rhIL-4; OPC;
 KW osteoprogenitor cell; odontoprogenitor cell; alveolar bone loss; RA;
 KW rheumatoid arthritis; periodontitis; osteolysis; bone metastasis;
 KW osteoparthritis; cytostatic; anticaries; gene-therapy; osteoblast;
 KW osteocalcin; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200160983-A2.
 XX
 PD 23-AUG-2001.
 XX
 XX 09-FEB-2001; 2001WO-US004274.
 XX
 XX 18-FEB-2000; 2000US-00507239.
 XX
 PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.
 XX
 PI Keeping HS, Reichner JS;
 XX
 DR WPI; 2001-514769/56.
 XX
 PT An isolated odontoprogenitor cell or osteoprogenitor cell for the
 PT treatment of rheumatoid arthritis both comprising a nucleic acid encoding
 PT an anti-inflammatory peptide.
 XX
 XX Example 3; Page 4; 28pp; English.
 XX
 CC The patent discloses compositions and methods to deliver an anti-
 CC inflammatory composition, e.g. recombinant human interleukin-4 (rhIL-4),
 CC to build (or rebuild) bone tissue. The patent also relates to
 CC osteoprogenitor cells (OPCs) or odontoprogenitor cells comprising a
 CC nucleic acid encoding an anti-inflammatory peptide. The odonto-
 CC progenitor cells are useful for treating a mammal at risk of developing
 CC periodontitis or is suffering from or is at risk of developing alveolar
 CC bone loss due to periodontal disease. OPCs are useful for treating a (RA),
 CC mammal suffering from or at risk of developing rheumatoid arthritis (RA),
 CC periapical or endochondral bone loss, artificial joint particle-induced

CC osteolysis, or osteolytic bone metastases. They are used to inhibit the
CC formation of osteoclasts which resorb bone and to stimulate the growth of
CC new bones. The OCs and odontoblast precursor cells are useful for inhibiting
CC osteolysis and for inducing differentiation of a bone marrow stromal
CC cells. These cells are also used in gene therapy. The present sequence is
CC human osteoblast-specific promoter DNA, osteocalcin promoter
XX

Sequence 597 BP; 107 A; 190 C; 187 G; 113 T; 0 U; 0 Other;

Query Match 85.2%; Score 23; DB 5; Length 597;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTAAGTACCGGGTGAACGGGG 23
DB 65 GGTAAGTACCGGGTGAACGGGG 87

RESULT 13

AD48137 standard; DNA; 2650 BP.

AC AD48137;

DT 24-FEB-2003 (first entry)

DE Human pleiotrophin/OSF-1 DNA.

XX Brain; tumour protein target; Tbt; ischaemic stroke; cancer; epilepsy;
XX schizophrenia; depression; Alzheimer's disease; Parkinson's disease;
XX Huntington's chorea; traumatic head injury; dementia; stupor; headache;
XX coma; vertigo; weakness; myasthenia gravis; cerebrovascular disorder;
XX infection; multiple sclerosis; pregnancy; medical illness; vasculitis;
XX metabolic deficiency; cerebroprotective; antidepressant; antibacterial;
XX cytosolic; neurotrophic; analgesic; fungicide; virucide; pleiotrophin;
XX OSF-1; human; gene; ds.

OS Homo sapiens.

Key Location/Qualifiers
FT CDS 1542..2048
FT /tag= a
FT /product= "Human pleiotrophin/OSF-1 protein"

XX WO200276510-A1.

PD 03-OCT-2002.

PF 22-MAR-2002; 2002WO-US008992.

XX 23-MAR-2001; 2001US-00816703.

PR 17-OCT-2001; 2001US-00983000.

PA (AGT-) AGY THERAPEUTICS INC.

PI Mueller S, Melcher T, Chin DJ;

XX WPI; 2003-029903/02.

DR P-PSDB; AAE30342.

XX Developing active agents that modulate the activity of a brain tumor
XX protein target gene or gene product for treating e.g. stroke or cancer,
XX comprises contacting an agent with a brain tumor protein.

Claim 1; Page 130-132; 135pp; English.

XX The invention relates to a method for developing biologically active
XX agents that modulate activity of a brain tumor protein target (Tbt) gene
XX or gene product. The method is useful for developing biologically active
XX agents that modulate the activity of a brain tumor protein target gene
XX or gene product. Compounds that bind to the brain tumor proteins are
XX useful for treating e.g. ischaemic stroke, brain cancer, epilepsy,
XX schizophrenia, depression, Alzheimer's disease, Parkinson's disease,

CC Huntington's chorea, traumatic head injury, dementia, stupor, headache,
CC coma, vertigo, weakness, myasthenia gravis, cerebrovascular disorders,
CC infectious disorders (including fungal, bacterial, viral and parasitic
CC infections), multiple sclerosis, and other complications associated with
CC pregnancy, medical illness, alcohol and substance abuse, toxins and
CC metabolic deficiencies. The brain tumor proteins may also be used to
CC raise antibodies. The present sequence is human pleiotrophin/OSF-1 DNA
XX used to illustrate the method of the invention

Sequence 2650 BP; 639 A; 657 C; 649 G; 705 T; 0 U; 0 Other;

Query Match 85.2%; Score 23; DB 8; Length 2650;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTAAGTACCGGGTGAACGGGG 23
DB 343 GGTAAGTACCGGGTGAACGGGG 365

RESULT 14

ADN06011 standard; cDNA; 2650 BP.

AC ADN06011;

DT 01-JUL-2004 (first entry)

DE Antipsoriatic cDNA sequence #1241.

XX ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX Homo sapiens.

XX WO2004028479-A2.

PD 08-APR-2004.

PF 25-SEP-2003; 2003WO-US030907.

PR 25-SEP-2002; 2002US-0414006P.

PA (GENT) GENENTECH INC.

PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;

XX WPI; 2004-305105/28.

DR P-PSDB; ADN06012.

XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.

Claim 1; SEQ ID NO 2406; 3069bp; English.

XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polynucleotides of the invention.

Sequence 2650 BP; 639 A; 657 C; 649 G; 705 T; 0 U; 0 Other;

Query Match 85.2%; Score 23; DB 12; Length 2650;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTAAGTACCGGGTGAACGGGG 23
DB 343 GGTAAGTACCGGGTGAACGGGG 365

RESULT 15

AB234742
ID AB234742 standard; cDNA; 2657 BP.

XX
AC AB234742;

XX
DT 04-FEB-2003 (first entry)

XX
DE Coding sequence SEQ ID 100, downregulated in osteogenesis.

XX
KM Osteoporosis; osteogenesis modulator; gene therapy; osteogenesis;
XX osteoporosis; bone disease; downregulator; human; ss.

XX
OS Homo sapiens.

XX
PN W0200281745-A2.

XX
PD 17-OCT-2002.

XX
PF 05-APR-2002; 2002WO-IB002211.

XX
PR 05-APR-2001; 2001US-0281400P.

XX
PA (AVET) AVENTIS PHARMA SA.

PI Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;

PI Connolly T, Jackson A, Bushnell SE, Rawadi G;

XX
DR WPI; 2003-058567/05.

XX
PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
XX useful for bone disease therapy in subject.

XX
PS Claim 27; Page 126-127; 237pp; English.

XX
CC The present invention relates to novel nucleotide sequences, which are
XX differentially expressed in models of osteogenesis upon being put in
XX contact with a stimulator of osteogenesis. The present sequence is one
XX such sequence. This sequence can be used for diagnosing osteoporosis/bone
XX disease in a patient, promoting osteogenesis and/or preventing
XX osteoporosis/bone disease. The present sequence encodes a secreted
XX protein

XX
SQ Sequence 2657 BP; 640 A; 656 C; 648 G; 706 T; 0 U; 7 Other;

Query Match 85.2%; Score 23; DB 8; Length 2657;

Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGGTGAACGGG 23
DB 343 GGTGACTCACCAGGTGAACGGG 365

Search completed: November 29, 2004, 14:21:32
Job time : 403 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 13:45:01 ; Search time 89 seconds

(without alignments)
215.633 Million cell updates/sec

Title: US-10-073-064-2

Perfect score: 27

Sequence: 1 GGTGACTCACCAGGCGGTGAACGGGGCATT 27

Scoring table:

IDENTITY NUC
Gapop 10⁻⁰, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/pdata/1/ina/5A.COMB.seq:*
2: /cgn2_6/pdata/1/ina/5B.COMB.seq:*
3: /cgn2_6/pdata/1/ina/6A.COMB.seq:*
4: /cgn2_6/pdata/1/ina/6B.COMB.seq:*
5: /cgn2_6/pdata/1/ina/PCITUS.COMB.seq:*
6: /cgn2_6/pdata/1/ina/backfiles1.seq:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	27	1	US-07-789-919A-2
2	27	100.0	27	1	US-08-209-846A-2
3	27	100.0	27	2	US-08-472-809B-2
4	27	100.0	27	3	US-08-438-265-2
5	23	85.2	301	5	PCT-US94-05659-2
6	18.2	67.4	1440	4	US-10-141-634-1
7	18	66.7	624	4	US-09-489-039A-1953
8	18	66.7	768	4	US-09-489-039A-1930
9	18	66.7	4245	3	US-09-276-531-16
10	18	66.7	4500	4	US-09-087-031E-27
11	18	66.7	4403765	3	US-09-103-840A-2
12	18	66.7	4411529	3	US-09-103-840A-1
13	17.6	65.2	2138	4	US-09-056-556-183
14	17.6	65.2	2138	4	US-09-072-596-178
15	17.6	65.2	2138	4	US-09-072-596-178
16	17.4	64.4	7194	4	US-09-601-326-76
17	17.4	64.4	15420	4	US-09-601-326-54
18	17	63.0	790	4	US-09-976-594-796
19	17	63.0	2933	3	US-09-262-773-209
20	17	63.0	23071	3	US-09-262-773-210
21	17	63.0	4403765	3	US-09-103-840A-2
22	17	63.0	4411529	3	US-09-103-840A-1
23	16.8	62.2	5917	4	US-09-780-175-17
24	16.6	61.5	321	4	US-09-401-064-348
25	16.6	61.5	14462	4	US-09-843-250-9
26	16.4	60.7	30	3	US-08-865-960-27
27	16.4	60.7	629	4	US-09-477-135A-123

C 28	16.4	60.7	630	4	US-09-489-039A-938	Sequence 938, App
C 29	16.4	60.7	1170	4	US-09-489-039A-936	Sequence 936, App
C 30	16.4	60.7	1629	4	US-09-489-039A-4586	Sequence 4586, App
C 31	16.2	60.0	3177	4	US-09-252-991A-12534	Sequence 12534, A
C 32	16	59.3	192	2	US-08-604-989A-6	Sequence 6, Appl
C 33	16	59.3	212	4	US-09-016-434-50	Sequence 50, Appl
C 34	16	59.3	243	4	US-09-489-039A-5679	Sequence 5679, Ap
C 35	16	59.3	259	4	US-09-513-999C-2402	Sequence 2402, Ap
C 36	16	59.3	295	4	US-09-919-172-44	Sequence 44, Appl
C 37	16	59.3	411	3	US-09-059-369-16	Sequence 16, Appl
C 38	16	59.3	596	4	US-08-961-527-347	Sequence 347, App
C 39	16	59.3	735	4	US-09-583-110-2085	Sequence 2085, Ap
C 40	16	59.3	1022	3	US-09-222-575-67	Sequence 67, Appl
C 41	16	59.3	1022	4	US-09-389-681-67	Sequence 67, Appl
C 42	16	59.3	1022	4	US-09-620-405B-67	Sequence 67, Appl
C 43	16	59.3	1022	4	US-09-339-338-67	Sequence 67, Appl
C 44	16	59.3	1022	4	US-09-433-826B-67	Sequence 67, Appl
C 45	16	59.3	1022	4	US-09-604-287A-67	Sequence 67, Appl

ALIGNMENTS

RESULT 1
US-07-789-919A-2
Sequence 2, Application US/0789919A
Patent No. 5298422
GENERAL INFORMATION:
APPLICANT: Schwartz, Robert J.
APPLICANT: Demayo, Franco
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: Myogenic Vector Systems
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thomas D. Paul
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/789,919A
FILING DATE: 19911106
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5379
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
TELEX: 762829
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-07-789-919A-2
Query Match 100.0%; Score 27; DB 1; Length 27;
Best local Similarity 100.0%; Pred. No. 0.0012;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 GGTGACTCACC GG GTGAACGGGGCATT 27

RESULT 2

US-08-209-846A-2
Sequence 2, Application US/08209846A

Patent No. 5756264

GENERAL INFORMATION:

APPLICANT: Schwartz, Robert J.

APPLICANT: Demayo, Franco

APPLICANT: O'Malley, Bert W.

TITLE OF INVENTION: Expression Vector Systems and Methods of

TITLE OF INVENTION: Use

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon

STREET: 611 West Sixth St.

CITY: Los Angeles

STATE: CA

COUNTRY: USA

ZIP: 90017

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/209,846A

FILING DATE: 09-JUN-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/789,919

FILING DATE: 06-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: Knight, Matthew W.

REGISTRATION NUMBER: 36,846

REFERENCE/DOCKET NUMBER: 204/302

TELECOMMUNICATION INFORMATION:

TELEPHONE: 213-489-1600

TELEFAX: 213-955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 27 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-209-846A-2

Query Match 100.0%; Score 27; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 GGTGACTCACC GG GTGAACGGGGCATT 27

RESULT 3

US-08-472-809B-2
Sequence 2, Application US/08472809B

Patent No. 5925564

GENERAL INFORMATION:

APPLICANT: Schwartz, Robert J.

APPLICANT: Demayo, Franco

APPLICANT: O'Malley, Bert W.

TITLE OF INVENTION: Expression Vector Systems and

TITLE OF INVENTION: Method of Use

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon

STREET: 633 West Fifth Street

STREET: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 MB

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,809B

FILING DATE: June 7, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/209,846

FILING DATE: March 9, 1994

APPLICATION NUMBER: 07/789,919

FILING DATE: No. 5925564ember 6, 1991

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 214/212

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 27 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-472-809B-2

Query Match 100.0%; Score 27; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 GGTGACTCACC GG GTGAACGGGGCATT 27

RESULT 4

US-08-438-265-2
Sequence 2, Application US/08438265

Patent No. 6361984

GENERAL INFORMATION:

APPLICANT: Schwartz, Robert J.

APPLICANT: Demayo, Franco

APPLICANT: O'Malley, Bert W.

TITLE OF INVENTION: Expression Vector Systems and Methods of

TITLE OF INVENTION: Use

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSER: Lyon & Lyon

STREET: 611 West Sixth St.

CITY: Los Angeles

STATE: CA

COUNTRY: USA

ZIP: 90017

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/438,265

FILING DATE: 09-May-1995

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/368,776
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/789,919
FILING DATE: 06-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Knight, Matthew W.
REGISTRATION NUMBER: 36,846
REFERENCE/DOCKET NUMBER: 204/302
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-489-1600
TELEFAX: 213-955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-438-265-2

Query Match
Best Local Similarity 100.0%; Score 27; DB 3; Length 27;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACC GGGTGAACGGGGCATT 27
Db 1 GGTGACTCACC GGGTGAACGGGGCATT 27

RESULT 5
PCT-US94-05659-2
Sequence 2, Application PC/TUS9405659
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: TNF RESPONSIVE ELEMENT, TNF-INDUCED DNA-BINDING
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Milflia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05659
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: FDC93-01 FF
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 301 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US94-05659-2

Query Match 85.2%; Score 23; DB 5; Length 301;

Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACC GGGTGAACGGGGCATT 23
Db 240 GGTGACTCACC GGGTGAACGGGGCATT 262

RESULT 6
US-10-141-634-1/c
Sequence 1, Application US/10141634
Patent No. 6734010
GENERAL INFORMATION:
APPLICANT: Yu, Xuechuan
APPLICANT: Xie, Qiongshu
APPLICANT: Aduin, Alejandro
APPLICANT: Walke, D. Wade
TITLE OF INVENTION: No. 6734010el Human Kinases and Polynucleotides Encoding the Sar
FILE REFERENCE: LEX-0344-USA
CURRENT APPLICATION NUMBER: US/10/141,634
CURRENT FILING DATE: 2002-05-08
PRIOR APPLICATION NUMBER: US 60/289,727
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1440
TYPE: DNA
ORGANISM: homo sapiens
US-10-141-634-1

Query Match 67.4%; Score 18.2; DB 4; Length 1440;
Best Local Similarity 87.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 ACTCACC GGGTGAACGGGGCATT 27
Db 812 ACCCTCGGGGTGAACGGGGCATT 790

RESULT 7
US-09-489-039A-1953
Sequence 1953, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 1953
LENGTH: 624
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1953

Query Match 66.7%; Score 18; DB 4; Length 624;
Best Local Similarity 80.8%; Pred. No. 26;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GTGACTCACC GGGTGAACGGGGCATT 27
Db 434 GTGACTCACC GGGTGAACGGGGCATT 459

RESULT 8
US-09-489-039A-1930/c
Sequence 1930, Application US/09489039A
Patent No. 6610836

GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO: 1930
LENGTH: 768
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1930

Query Match 66.7%; Score 18; DB 4; Length 768;
Best Local Similarity 80.8%; Pred. No. 27;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GTGACTCACCAGGTTGACCGGGGCACT 27
Db 291 GTGACTCACCAGGTTGACCGGGGCACT 266

RESULT 9
US-09-276-531-16/c
Sequence 16, Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA: US/09/276,531
APPLICATION NUMBER: US/09/276,531
FILING DATE: Herewith
CLASSIFICATION:
APPLICATION NUMBER: 60/079,677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murty, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 4245 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: BRAITUT21
CLONE: 2526432
US-09-276-531-16

Query Match 66.7%; Score 18; DB 3; Length 4245;
Best Local Similarity 80.8%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GTGACTCACCAGGTTGACCGGGGCACT 27
Db 1045 GTGACTCACCAGGTTGACCGGGGCACT 1020

RESULT 10
US-09-087-031E-27/c
Sequence 27, Application US/09087031E
Patent No. 6479255
GENERAL INFORMATION:
APPLICANT: Rubin, Jeffrey S.
APPLICANT: Finch, Paul
APPLICANT: Aaronson, Stuart
APPLICANT: He, Xi
TITLE OF INVENTION: HUMAN FRP AND FRAGMENTS THEREOF INCLUDING METHODS FOR USING THEM
FILE REFERENCE: 11613.13US11
CURRENT APPLICATION NUMBER: US/09/087,031E
CURRENT FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 09/087,031
PRIOR FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 60/050,417
PRIOR FILING DATE: 1997-06-23
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 27
TYPE: DNA
LENGTH: 4500
ORGANISM: Homo sapiens
US-09-087-031E-27

Query Match 66.7%; Score 18; DB 4; Length 4500;
Best Local Similarity 80.8%; Pred. No. 34;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GTGACTCACCAGGTTGACCGGGGCACT 27
Db 1359 GTGACTCACCAGGTTGACCGGGGCACT 1334

RESULT 11
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 66.7%; Score 18; DB 3; Length 4403765;
Best Local Similarity 80.8%; Pred. No. 63;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GTGACTCAGCGGGTGACGGGGCATT 27
DB 2204211 GTGTCCACCGCGGTGACCGGCATT 2204186

RESULT 12

US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24386-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 66.7%; Score 18; DB 3; Length 4411529;
Best Local Similarity 80.8%; Pred. No. 63;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GTGACTCAGCGGGTGACGGGGCATT 27
DB 2206912 GTGTCCACCGCGGTGACCGGCATT 2206887

RESULT 13

US-09-056-556-183
Sequence 183, Application US/09056556
Patent No. 6350456
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-Apr-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 183:
SEQUENCE CHARACTERISTICS:
LENGTH: 2138 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-056-556-183

Query Match 65.2%; Score 17.6; DB 3; Length 2138;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGACTCAGCGGGTGACGGGGC 24
DB 1798 GTGTCTCAGCGGGTGACGAGGC 1821

RESULT 14

US-09-072-596-178
Sequence 178, Application US/09072596
Patent No. 6458366
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, David C.
APPLICANT: Campos-Neto, Antonia
APPLICANT: Houghton, Raymond
APPLICANT: Veckick, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 350
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-May-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 178:
SEQUENCE CHARACTERISTICS:
LENGTH: 2138 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-072-596-178

Query Match 65.2%; Score 17.6; DB 4; Length 2138;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGACTCAGCGGGTGACGGGGC 24
DB 1798 GTGTCTCAGCGGGTGACGAGGC 1821

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 14:14:46 ; Search time 4972 Seconds

(without alignments)
29.430 Million cell updates/sec

Title: US-10-073-064-2

Perfect score: 27
Sequence: 1 GGTACTCACCAGCGGTGACGGGGCATT 27

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3640242 seqs, 2709731945 residues

Total number of hits satisfying chosen parameters: 7280484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/FCI_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/FCIUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10F_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	27	US-10-073-064-2	Sequence 2, Appl1
2	23	85.2	597	US-10-683-516-3	Sequence 3, Appl1
3	23	85.2	2657	US-10-171-311-193	Sequence 133, App
4	18.2	67.4	132	US-10-029-386-19439	Sequence 15439, A
5	18.2	67.4	391	US-10-027-632-26905	Sequence 26905, A
6	18.2	67.4	391	US-10-027-632-26905	Sequence 26905, A
7	18.2	67.4	545	US-10-029-386-5683	Sequence 5683, App
8	18.2	67.4	774	US-10-165-800-23	Sequence 23, Appl1
9	18.2	67.4	995	US-10-311-034-51	Sequence 51, Appl1
10	18.2	67.4	1223	US-09-764-868-257	Sequence 257, Appl1
11	18.2	67.4	1440	US-10-141-634-1	Sequence 1, Appl1
12	18.2	67.4	1452	US-10-165-800-21	Sequence 21, Appl1

Result No.	Score	Query Match	Length	ID	Description
13	18.2	67.4	1579	US-10-094-749-1381	Sequence 1381, App
14	18.2	67.4	1581	US-10-478-146-17	Sequence 17, Appl1
15	18.2	67.4	1619	US-10-104-047-1657	Sequence 1657, App
16	18.2	67.4	1778	US-10-098-841-257	Sequence 257, App
17	18.2	67.4	5340	US-10-311-455-207	Sequence 207, App
18	18.2	67.4	1063	US-10-283-122A-22929	Sequence 22929, A
19	18	66.7	2602	US-10-666-851-1	Sequence 1, Appl1
20	18	66.7	3215	US-09-925-301-51	Sequence 51, Appl1
21	18	66.7	4452	US-09-874-298-28	Sequence 28, Appl1
22	18	66.7	4500	US-10-138-434A-27	Sequence 27, Appl1
23	18	66.7	21230	US-10-087-192-601	Sequence 601, Appl1
24	17.6	65.2	1075	US-10-425-114-1745	Sequence 1745, App
25	17.6	65.2	1102	US-10-425-114-16946	Sequence 16946, App
26	17.6	65.2	1518	US-10-739-930-3057	Sequence 3057, App
27	17.6	65.2	2138	US-10-193-002-178	Sequence 178, App
28	17.6	65.2	2138	US-10-084-843-183	Sequence 183, App
29	17.6	65.2	2967	US-09-938-842A-432	Sequence 432, App
30	17.6	65.2	2967	US-09-938-842A-432	Sequence 432, App
31	17.6	65.2	3323	US-10-225-066A-151	Sequence 151, App
32	17.6	65.2	3323	US-10-374-780A-237	Sequence 237, App
33	17.6	65.2	16181	US-09-764-847-1426	Sequence 1426, App
34	17.6	65.2	16181	US-09-764-847-1426	Sequence 1426, App
35	17.6	65.2	16181	US-10-092-154-1826	Sequence 1826, App
36	17.6	65.2	495269	US-10-398-221-8	Sequence 8, Appl1
37	17.6	65.2	3011208	US-10-398-221-8	Sequence 2058, App
38	17.4	64.4	686	US-10-424-559-110292	Sequence 110292, App
39	17.4	64.4	864	US-10-369-493-24087	Sequence 24087, A
40	17.4	64.4	947	US-10-027-632-161355	Sequence 161355, A
41	17.4	64.4	947	US-10-027-632-161355	Sequence 161355, A
42	17.4	64.4	947	US-10-027-632-161355	Sequence 161355, A
43	17.4	64.4	947	US-10-027-632-161355	Sequence 161355, A
44	17.4	64.4	1587	US-10-260-046-15	Sequence 15, Appl1
45	17.4	64.4	1611	US-10-260-046-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-10-073-064-2
Sequence 2, Appl1
Publication No. US20030104407A1
GENERAL INFORMATION:
APPLICANT: Schwartz, Robert J.
O'Malley, Franco
Demayo, Berto W.
TITLE OF INVENTION: Expression Vector Systems and Methods of Use
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 611 West Sixth St.
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/073.064
FILING DATE: 12-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/438.265
FILING DATE: 09-May-1995
APPLICATION NUMBER: 08/368.776
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/789.919
FILING DATE: 06-NOV-1991
ATTORNEY/AGENT INFORMATION:

```

; NAME: Knight, Matthew W.
; REGISTRATION NUMBER: 36, 846
; REFERENCE/DOCKET NUMBER: 204/302
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-489-1600
; TELEFAX: 213-955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-073-064-2

Query Match      100.0%; Score 27; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGGCGGTGAACGGGCATT 27
DB 1 GGTGACTCACCAGGCGGTGAACGGGCATT 27

RESULT 2
US-10-683-516-3
; Sequence 3, Application US/10683516
; Publication No. US20040126364A1
; GENERAL INFORMATION:
; APPLICANT: Keeping, Hugh S
; APPLICANT: Reicher, Jonathan S
; TITLE OF INVENTION: Treatment for Bone Disorders
; FILE REFERENCE: 21486-028CON
; CURRENT APPLICATION NUMBER: US/10/683,516
; PRIOR FILING DATE: 2003-10-09
; PRIOR APPLICATION NUMBER: 09/507,239
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-683-516-3

Query Match      85.2%; Score 23; DB 17; Length 597;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGGCGGTGAACGGGC 23
DB 65 GGTGACTCACCAGGCGGTGAACGGGC 87

RESULT 3
US-10-171-311-193
; Sequence 193, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavardu, Manjula
; APPLICANT: Hoersht, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
```

```

; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 193
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 2623..2624, 2625, 2626, 2627, 2628, 2629
; OTHER INFORMATION: n = A,T,C or G
US-10-171-311-193

Query Match      85.2%; Score 23; DB 14; Length 2657;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGGCGGTGAACGGGC 23
DB 343 GGTGACTCACCAGGCGGTGAACGGGC 365

RESULT 4
US-10-029-386-19439/C
; Sequence 19439, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Harzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEW07CA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19439
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000392.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.51
; OTHER INFORMATION: SWISSPROT HIT: Q9RSK7, EVALUE 7.90e-02
; OTHER INFORMATION: NT HIT: AL603647.1, EVALUE 4.40e-01
; OTHER INFORMATION: EST_HUMAN HIT: AW069362.1, EVALUE 7.00e-38
US-10-029-386-19439

Query Match      67.4%; Score 18.2; DB 15; Length 132;
Best Local Similarity 87.0%; Pred. No. 82;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ACTCACCAGGCGGTGAACGGGCATT 27
DB 55 ACCCTCGGGGTGAACGGGCATT 33

RESULT 5
US-10-027-632-26905
; Sequence 26905, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```



```

? TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
? TITLE OF INVENTION: Polymorphisms in the Human Genome
? FILE REFERENCE: 109827.129
? CURRENT APPLICATION NUMBER: US/10/027,632
? CURRENT FILING DATE: 2002-04-30
? PRIOR APPLICATION NUMBER: US 60/218,006
? PRIOR FILING DATE: 2000-07-12
? PRIOR APPLICATION NUMBER: US 60/198,676
? PRIOR FILING DATE: 2000-04-20
? PRIOR APPLICATION NUMBER: US 60/193,483
? PRIOR FILING DATE: 2000-03-29
? PRIOR APPLICATION NUMBER: US 60/185,218
? PRIOR FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/167,363
? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/156,358
? PRIOR FILING DATE: 1999-09-28
? PRIOR APPLICATION NUMBER: US 60/146,002
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 325720
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 26905
? LENGTH: 391
? TYPE: DNA
? ORGANISM: Human
? US-10-027-632-26905

```

Query Match	67.4%	Score 18.2	DB 13	Length 391
Best Local Similarity	87.0%	Pred. No. 81		
Matches	20	Conservative	0	Mismatches 3
				Indels 0
				Gaps 0

QY	3	TCACTACCGG	GTGAACGGGGCA	25
DB	88	TCAGTCACCGG	AGGAACGGGGCA	110

```

RESULT' 6
US-10-027-632-26905
; Sequence 26905, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 109827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26905
; LENGTH: 391
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-26905

Query Match      67.4%; Score 18.2; DB 15; Length 391;
Best Local Similarity 87.0%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0

```

```

QY      3  TGACTACCCGGGTGAACGGGGCA  25
      ||| ||| ||| ||| ||| |||
Db      88  TGACTACCCGGGAGAACCTGGCA  110

```

```

RESULT 7
US-10-029-386-5683/c
; Sequence 5683, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5683
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000392.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.51
; OTHER INFORMATION: EST HUMAN HIT: AW069362.1, EVALUE 8.00e-38
; OTHER INFORMATION: SWISSPROT HIT: Q9V8U5, EVALUE 4.00e-02
US-10-029-386-5683

```

Query Match	67.4%	Score 18.2;	DB 15,	Length 545;
Best Local Similarity	87.0%	Pred. No. 81;		
Matches 20;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0

```

OY      5  ACTCACGGGTGAACGGGGCATT  27
Db      266  ACCCTCGGGGTGAACGGGGCATT  244

RESULT 8
US-10-165-800-23/c
; Sequence 23, Application US/10165800
; Publication No. US20030092116A1
; GENERAL INFORMATION:
; APPLICANT: Chun, Miyoung
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: No. US20030092116A1 Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Fragmentation Factor-Like, Phospholipid Scramblase-Like, DNA
; TITLE OF INVENTION: and Arpase-Like Molecules and Uses Therefor
; FILE REFERENCE: 35800/247838
; CURRENT APPLICATION NUMBER: US/10/165,800
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 09/790,179
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/185,609
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/790,838
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/185,946
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/790,180
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/185,947
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/795,038
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/186,234

```

```

; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/781,677
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/181,705
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-165-800-23

Query Match      67.4%; Score 18.2; DB 14; Length 774;
Best Local Similarity 87.0%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ACTCACC GG GTGAACGGGCATT 27
DB 596 ACCCTCGGGGTGAACGGGCATT 574

RESULT 9
US-10-311-034-51/c
; Sequence 51, Application US/10311034
; Publication No. US20040023242A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: BANDMAN, Olga
; APPLICANT: BOROMSKY, Mark L.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyang, Alma M.
; APPLICANT: GREENWALD, Sara R.
; APPLICANT: RAMKUTVAR, Jayalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: KEARNEY, Liam
; APPLICANT: BURFORD, Neil
; APPLICANT: NGUYEN, Daniel B.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HE, Ann
; APPLICANT: THORNTON, Michael
; APPLICANT: HAPALIA, April
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: LO, Terence P.
; APPLICANT: KHAH, Parrah A.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: DING, Li
; APPLICANT: GREYER, Megan
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: BATRA, Sajeev
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0125 PCT
; CURRENT APPLICATION NUMBER: US/10/311,034
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,073; 60/213,467; 60/215,651; 60/216,605; 60/218,372;
; 60/228,056
; PRIOR FILING DATE: 2000-06-15; 2000-06-23; 2000-06-30; 2000-07-07; 2000-07-13; 2000-0
25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PERL Program
```

```

; SEQ ID NO 51
; LENGTH: 995
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023242A1 621293CBI
US-10-311-034-51

Query Match      67.4%; Score 18.2; DB 16; Length 995;
Best Local Similarity 87.0%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ACTCACC GG GTGAACGGGCATT 27
DB 853 ACCCTCGGGGTGAACGGGCATT 831

RESULT 11
US-10-141-634-1/c

Query Match      67.4%; Score 18.2; DB 9; Length 1223;
Best Local Similarity 87.0%; Pred. No. 81;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ACTCACC GG GTGAACGGGCATT 27
DB 853 ACCCTCGGGGTGAACGGGCATT 831
```

```
; Sequence 1, Application US/10141634
; Publication No. US20030008365A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Xie, Qiongshu
; APPLICANT: Abuin, Alejandro
; APPLICANT: Walke, D. Wade
; TITLE OF INVENTION: No. US20030008365A1 Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0344-USA
; CURRENT APPLICATION NUMBER: US/10/141,634
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: US 60/289,727
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1440
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-141-634-1

Query Match      67.4%; Score 18.2; DB 14; Length 1440;
Best Local Similarity 87.0%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      5 ACTCACC GG GTGAACGGGCATT 27
Db      812 ACCCTGGGGGTGAACGGGCATT 790

RESULT 12
US-10-165-800-21/c
; Sequence 21, Application US/10165800
; Publication No. US20030092116A1
; GENERAL INFORMATION:
; APPLICANT: Chun, Mlyoung
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Kapeller-Liberman, Rosana
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: No. US20030092116A1 Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Adenylate Kinase, Phospholipid Scramblase-Like, DNA
; TITLE OF INVENTION: Fragmentation Factor-Like, Phosphatidylserine Synthase-Like,
; FILE REFERENCE: 35800/247838
; CURRENT APPLICATION NUMBER: US/10/165,800
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 09/790,179
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/185,609
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/790,838
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/185,946
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/790,180
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/185,947
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/795,038
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/186,234
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/781,677
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/181,705
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (219)...(995)
US-10-165-800-21

Query Match      67.4%; Score 18.2; DB 14; Length 1452;
Best Local Similarity 87.0%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      5 ACTCACC GG GTGAACGGGCATT 27
Db      814 ACCCTGGGGGTGAACGGGCATT 792

RESULT 13
US-10-094-749-1381/c
; Sequence 1381, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TEISUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1381
; LENGTH: 1579
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-1381

Query Match      67.4%; Score 18.2; DB 15; Length 1579;
Best Local Similarity 87.0%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      5 ACTCACC GG GTGAACGGGCATT 27
Db      841 ACCCTGGGGGTGAACGGGCATT 819

RESULT 14
US-10-478-146-17/c
; Sequence 17, Application US/10478146
; Publication No. US20040203097A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; YUE, Henry;
; APPLICANT: LU, Dyrng Aina M.; AZIMZAI, Yalda;
; APPLICANT: DING, Li JIE; Ernestine A.;
; APPLICANT: HANFALTA, April J.A.; BECHA, Shanya D.;
; APPLICANT: TANG, Y. Tom; LAU, Preeti G.;
; APPLICANT: GRIFFIN, Jennifer A.; GURURAJAN, Rajagopal;
; APPLICANT: RAMKUMAR, Jayalaxmi; ELIOTT, Vicki S.;
```

APPLICANT: ARVIZU, Chandra S.; LUO, Wen;
APPLICANT: SWARNAKAR, Anita; DUGGAN, Brendan M.;
APPLICANT: TRAN, Uyen K.; CHAMLA, Narinder K.;
APPLICANT: GANDHI, Ameena R.; YAO, Monique G.;
APPLICANT: KHAN, Farrah A.; BAUGHN, Nathan R.;
APPLICANT: BOROWSKY, Mark L.; ZEBARJADIAN, Yeganeh;
APPLICANT: RICHARDSON, Thomas W.; MARQUIS, Joseph P.;
APPLICANT: CHEN, David; JIN, Pei
TITLE OF INVENTION: KINASES AND PHOSPHATASES
FILE REFERENCE: PF-0995 USN
CURRENT APPLICATION NUMBER: US/10/478,146
PRIOR FILING DATE: 2003-11-18
PRIOR APPLICATION NUMBER: PCT/US02/16634
PRIOR FILING DATE: 2002-05-23
PRIOR APPLICATION NUMBER: US 60/293,665
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 60/298,712
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/303,418
PRIOR FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US 60/306,967
PRIOR FILING DATE: 2001-07-19
PRIOR APPLICATION NUMBER: US 60/308,183
PRIOR FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: US 60/343,007
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: US 60/357,675
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US 60/376,988
PRIOR FILING DATE: 2002-04-30
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PERL Program
SEQ ID NO 17
LENGTH: 1581
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: 72063274CBI
US-10-478-146-17

Query Match 67.4%; Score 18.2; DB 18; Length 1581;
Best Local Similarity 87.0%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ACTACCGGGGTGAACGGGCATT 27
Db 826 ACCCTCGGGGTGAACGGGCATT 804

RESULT 15
US-10-104-047-1657/c
Sequence 1657, Application US/10104047
Publication No. US20030236392A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: NO. US20030236392A1 full length cDNA
FILE REFERENCE: H1-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
PRIOR FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1657
LENGTH: 1619
TYPE: DNA
ORGANISM: Homo sapiens
US-10-104-047-1657

Query Match 67.4%; Score 18.2; DB 15; Length 1619;
Best Local Similarity 87.0%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ACTACCGGGGTGAACGGGCATT 27
Db 883 ACCCTCGGGGTGAACGGGCATT 861

Search completed: November 29, 2004, 17:46:30
Job time : 4974 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 13:45:02 ; Search time 3814 Seconds
(without alignments)
287.067 Million cell updates/sec

Title: US-10-073-064-2

Perfect score: 27
Sequence: 1 GGTACTCACCAGGAGTGAACGGAGGCAATT 27

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 44308572 seqs, 20275418765 residues

Total number of hits satisfying chosen parameters: 88617144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA Main:*

```
1: /cgn2_6/ptodata/1/pna/PCRTU51_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/PCRTU52_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US0923_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US109D_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US109E_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US109F_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US109G_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US109H_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US109I_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US109J_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US109K_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US109L_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US109M_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US109N_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US109O_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US109P_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US109Q_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US109R_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US109S_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US109T_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US109U_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US109V_COMB.seq:*
86: /cgn2_6/ptodata/1/pna/US109W_COMB.seq:*
87: /cgn2_6/ptodata/1/pna/US109X_COMB.seq:*
88: /cgn2_6/ptodata/1/pna/US109Y_COMB.seq:*
89: /cgn2_6/ptodata/1/pna/US109Z_COMB.seq:*
90: /cgn2_6/ptodata/1/pna/US109AA_COMB.seq:*
91: /cgn2_6/ptodata/1/pna/US109AB_COMB.seq:*
92: /cgn2_6/ptodata/1/pna/US109AC_COMB.seq:*
93: /cgn2_6/ptodata/1/pna/US109AD_COMB.seq:*
94: /cgn2_6/ptodata/1/pna/US109AE_COMB.seq:*
95: /cgn2_6/ptodata/1/pna/US109AF_COMB.seq:*
96: /cgn2_6/ptodata/1/pna/US109AG_COMB.seq:*
97: /cgn2_6/ptodata/1/pna/US109AH_COMB.seq:*
98: /cgn2_6/ptodata/1/pna/US109AI_COMB.seq:*
99: /cgn2_6/ptodata/1/pna/US109AJ_COMB.seq:*
100: /cgn2_6/ptodata/1/pna/US109AK_COMB.seq:*
101: /cgn2_6/ptodata/1/pna/US109AL_COMB.seq:*
102: /cgn2_6/ptodata/1/pna/US109AM_COMB.seq:*
103: /cgn2_6/ptodata/1/pna/US109AN_COMB.seq:*
104: /cgn2_6/ptodata/1/pna/US109AO_COMB.seq:*
105: /cgn2_6/ptodata/1/pna/US109AP_COMB.seq:*
106: /cgn2_6/ptodata/1/pna/US109AQ_COMB.seq:*
107: /cgn2_6/ptodata/1/pna/US109AR_COMB.seq:*
108: /cgn2_6/ptodata/1/pna/US109AS_COMB.seq:*
109: /cgn2_6/ptodata/1/pna/US109AT_COMB.seq:*
110: /cgn2_6/ptodata/1/pna/US109AU_COMB.seq:*
111: /cgn2_6/ptodata/1/pna/US109AV_COMB.seq:*
112: /cgn2_6/ptodata/1/pna/US109AW_COMB.seq:*
113: /cgn2_6/ptodata/1/pna/US109AX_COMB.seq:*
114: /cgn2_6/ptodata/1/pna/US109AY_COMB.seq:*
115: /cgn2_6/ptodata/1/pna/US109AZ_COMB.seq:*
116: /cgn2_6/ptodata/1/pna/US109BA_COMB.seq:*
```

117: /cgn2_6/prodata/1/pna/US6051_COMB.seq:*
118: /cgn2_6/prodata/1/pna/US6052_COMB.seq:*
119: /cgn2_6/prodata/1/pna/US6053_COMB.seq:*
120: /cgn2_6/prodata/1/pna/US6054_COMB.seq:*
121: /cgn2_6/prodata/1/pna/US6055_COMB.seq:*
122: /cgn2_6/prodata/1/pna/US6056_COMB.seq:*
123: /cgn2_6/prodata/1/pna/US6057_COMB.seq:*
124: /cgn2_6/prodata/1/pna/US6058_COMB.seq:*
125: /cgn2_6/prodata/1/pna/US6059_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	27	1 PCT-US92-09353-2	Sequence 2, Appl
2	27	100.0	27	US-08-209-846B-2	Sequence 2, Appl
3	27	100.0	27	US-08-451-399A-2	Sequence 2, Appl
4	27	100.0	27	US-08-451-882-2	Sequence 2, Appl
5	27	100.0	27	US-09-181-419A-2	Sequence 2, Appl
6	27	100.0	27	US-10-073-064-2	Sequence 2, Appl
7	23	85.2	57	US-09-400-967-6	Sequence 6, Appl
8	23	85.2	57	US-09-400-967-6	Sequence 6, Appl
9	23	85.2	301	US-08-066-372-2	Sequence 2, Appl
10	23	85.2	422	US-09-528-409-59789	Sequence 59789, A
11	23	85.2	422	US-09-933-524-59789	Sequence 59789, A
12	23	85.2	422	US-09-933-524-59789	Sequence 59789, A
13	23	85.2	422	US-09-933-524-59789	Sequence 59789, A
14	23	85.2	565	PCT-US02-22216-2	Sequence 2, Appl
15	23	85.2	565	PCT-US02-22216-2	Sequence 2, Appl
16	23	85.2	597	US-09-507-239-3	Sequence 3, Appl
17	23	85.2	597	US-10-683-516-3	Sequence 3, Appl
18	23	85.2	1604	US-10-170-235-4174	Sequence 4174, Ap
19	23	85.2	1604	US-10-170-235-4174	Sequence 2065, Ap
20	23	85.2	1604	US-06-453-050-1280	Sequence 1280, Ap
21	23	85.2	1604	US-06-453-135-1290	Sequence 1280, Ap
22	23	85.2	1604	US-06-453-135-1290	Sequence 1280, Ap
23	23	85.2	1667	US-10-170-235-4178	Sequence 4178, Ap
24	23	85.2	1667	US-06-452-680-2059	Sequence 2059, Ap
25	23	85.2	1667	US-06-453-050-1284	Sequence 1284, Ap
26	23	85.2	1667	US-06-453-135-1284	Sequence 1284, Ap
27	23	85.2	1765	US-09-724-676-14621	Sequence 14621, A
28	23	85.2	1765	US-09-724-676-14621	Sequence 14621, A
29	23	85.2	2571	US-09-724-676-14620	Sequence 14620, A
30	23	85.2	2571	US-09-724-676-14620	Sequence 14620, A
31	23	85.2	2650	PCT-US02-08992-16	Sequence 16, Appl
32	23	85.2	2650	US-10-821-801-119	Sequence 119, Appl
33	23	85.2	2657	PCT-US02-18638A-193	Sequence 193, Appl
34	23	85.2	2657	US-10-473-974-1193	Sequence 193, Appl
35	23	85.2	2657	US-10-473-974-1193	Sequence 100, Appl
36	23	85.2	2677	US-09-699-997-11736	Sequence 11736, A
37	23	85.2	2677	US-09-699-997-11736	Sequence 11736, A
38	23	85.2	2677	US-09-699-997-11736	Sequence 11736, A
39	23	85.2	2677	US-09-716-472-4937	Sequence 4937, Ap
40	23	85.2	2677	US-09-716-472-4937	Sequence 4937, Ap
41	23	85.2	2696	US-09-724-676-14629	Sequence 14629, A
42	23	85.2	2696	US-09-724-676-14629	Sequence 14629, A
43	23	85.2	2732	US-09-205-070-10104	Sequence 10104, A
44	23	85.2	2732	US-09-340-623-10104	Sequence 10104, A
45	23	85.2	2732	US-09-898-888-10104	Sequence 10104, A

ALIGNMENTS

RESULT 1
PCT-US92-09353-2
Sequence 2, Application PC/TUS9209353
GENERAL INFORMATION:

APPLICANT: Schwartz, Robert J.
APPLICANT: Demayo, Franco
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: Myogenic Vector Systems
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thomas D. Paul
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09353
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5379
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
TELEX: 762829
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOCHROMIC: NO
ANTI-SENSE: NO
PCT-US92-09353-2

Query Match 100.0%; Score 27; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTCACCGCGGTGAACGGGCATT 27
DB 1 GGTCACCGCGGTGAACGGGCATT 27

RESULT 2
US-08-209-846B-2
Sequence 2, Application US/08209846B
GENERAL INFORMATION:
APPLICANT: Schwartz, Robert J.
APPLICANT: Demayo, Franco
APPLICANT: O'Malley, Bert W.
TITLE OF INVENTION: EXPRESSION VECTOR SYSTEMS AND METHODS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/209,846B
; FILING DATE: March 9, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/789,919
; FILING DATE: November 6, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 204/302
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
;
; TELETYPE: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-209-846B-2
;
Query Match          100.0%; Score 27; DB 7; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTACCGGGTGAACGGGGCATT 27
Db 1 GGTGACTACCGGGTGAACGGGGCATT 27

RESULT 3
; US-08-451-399A-2
; Sequence 2, Application US/08451399A
; GENERAL INFORMATION:
; APPLICANT: Schwartz, Robert J.
; APPLICANT: Demayo, Franco
; APPLICANT: O'Malley, Bert W.
; TITLE OF INVENTION: EXPRESSION VECTOR SYSTEMS
; NUMBER OF SEQUENCES: 6
; NUMBER OF INVENTIONS: AND METHOD OF USE
; CURRENT APPLICATION DATA:
; ADDRESS: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM MS-DOS (version 6.22)
; SOFTWARE: Wordperfect (version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,399A
; FILING DATE: May 26, 1995
; CLASSIFICATION: 860
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,846
; FILING DATE: March 9, 1994
; APPLICATION NUMBER: 07/789,919
; FILING DATE: November 6, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkman, Charles S.
; REGISTRATION NUMBER: 38,077
; REFERENCE/DOCKET NUMBER: 212/136
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 955-0440
; TELEFAX: (213) 955-0440
;
; TELEX: 67-3510
;

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
; US-08-451-399A-2
;
Query Match          100.0%; Score 27; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTACCGGGTGAACGGGGCATT 27
Db 1 GGTGACTACCGGGTGAACGGGGCATT 27

RESULT 4
; US-08-451-882-2
; Sequence 2, Application US/08451882
; GENERAL INFORMATION:
; APPLICANT: Schwartz, Robert J.
; APPLICANT: Demayo, Franco
; APPLICANT: O'Malley, Bert W.
; TITLE OF INVENTION: Expression Vector Systems and Methods of
; NUMBER OF SEQUENCES: 6
; NUMBER OF INVENTIONS: Use
; CURRENT APPLICATION DATA:
; ADDRESS: Lyon & Lyon
; STREET: 611 West Sixth St.
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,882
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/209,846
; FILING DATE: 09-MAR-1994
; APPLICATION NUMBER: US 07/789,919
; FILING DATE: 06-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Knight, Matthew M.
; REGISTRATION NUMBER: 36,846
; REFERENCE/DOCKET NUMBER: 204/302
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-489-1600
; TELEFAX: 213-955-0440
; TELETYPE: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
; US-08-451-882-2
;
Query Match          100.0%; Score 27; DB 9; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTACCGGGTGAACGGGGCATT 27
Db 1 GGTGACTACCGGGTGAACGGGGCATT 27

```

```
RESULT 5
US-09-181-419A-2
; Sequence 2, Application US/09181419A
; GENERAL INFORMATION:
; APPLICANT: Valentis, Inc.
; APPLICANT: Schwartz, Robert J.
; APPLICANT: Demayo, Franco J.
; APPLICANT: O'Malley, Bert W.
; TITLE OF INVENTION: EXPRESSION VECTOR SYSTEMS AND METHOD OF USE
; FILE REFERENCE: SB Meyer: Valentis 236/277 US
; CURRENT APPLICATION NUMBER: US/09/181,419A
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 08/209,846
; PRIOR FILING DATE: 1994-03-09
; PRIOR APPLICATION NUMBER: 07/789,919
; PRIOR FILING DATE: 1991-11-06
; PRIOR APPLICATION NUMBER: 08/472,809
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 27
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-181-419A-2

Query Match
Best Local Similarity 100.0%; Score 27; DB 16; Length 27;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GGTGACTCAGCGGGTGAACGGGCATT 27
DB 1 GGTGACTCAGCGGGTGAACGGGCATT 27

RESULT 6
US-10-073-064-2
; Sequence 2, Application US/10073064
; GENERAL INFORMATION:
; APPLICANT: Schwartz, Robert J.
; APPLICANT: Demayo, Franco
; APPLICANT: O'Malley, Bert W.
; TITLE OF INVENTION: Expression Vector Systems and Methods of
; Use
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 611 West Sixth St.
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/073,064
; FILING DATE: 12-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,265
; FILING DATE: 09-May-1995
; APPLICATION NUMBER: 08/368,776
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/789,915
; FILING DATE: 06-Nov-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Knight, Matthew W.
; REGISTRATION NUMBER: 36,846
```

```
REFERENCE/DOCKET NUMBER: 204/302
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-489-1600
; TELEFAX: 213-955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-073-064-2

Query Match
Best Local Similarity 100.0%; Score 27; DB 46; Length 27;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GGTGACTCAGCGGGTGAACGGGCATT 27
DB 1 GGTGACTCAGCGGGTGAACGGGCATT 27

RESULT 7
US-09-400-967-6
; Sequence 5, Application US/09400967
; GENERAL INFORMATION:
; APPLICANT: Adams, John S.
; TITLE OF INVENTION: Nucleic Acid Encoding Vitamin D Response Element
; TITLE OF INVENTION: Binding Proteins, Products Related Thereto, and Methods
; FILE REFERENCE: P-CE 3489
; CURRENT APPLICATION NUMBER: US/09/400,967
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 57
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-400-967-6

Query Match
Best Local Similarity 85.2%; Score 23; DB 21; Length 57;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GGTGACTCAGCGGGTGAACGGGG 23
DB 3 GGTGACTCAGCGGGTGAACGGGG 25

RESULT 8
US-09-400-967-6/c
; Sequence 5, Application US/09400967
; GENERAL INFORMATION:
; APPLICANT: Adams, John S.
; TITLE OF INVENTION: Nucleic Acid Encoding Vitamin D Response Element
; TITLE OF INVENTION: Binding Proteins, Products Related Thereto, and Methods
; FILE REFERENCE: P-CE 3489
; CURRENT APPLICATION NUMBER: US/09/400,967
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 6
; LENGTH: 57
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-400-967-6

Query Match
Best Local Similarity 85.2%; Score 23; DB 21; Length 57;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
|||||
Db 56 GGTGACTCACC GG GTGAACGGGG 34

RESULT 9
US-08-066-372-2
Sequence 2, Application US/08066372

GENERAL INFORMATION:
APPLICANT: Li, Yi-Ping
APPLICANT: Stashenko, Philip
TITLE OF INVENTION: TNF-RESPONSIVE ELEMENT, TNF-INDUCED DNA-BINDING
TITLE OF INVENTION: PROTEIN AND METHOD OF INHIBITING TNF DOWN-REGULATION OF GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/066,372
FILING DATE: 19930520
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: FDC93-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 301 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-066-372-2

Query Match 85.2%; Score 23; DB 5; Length 301;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
|||||
Db 240 GGTGACTCACC GG GTGAACGGGG 262

RESULT 10
US-09-528-409-59789
Sequence 59789, Application US/09528409

GENERAL INFORMATION:
APPLICANT: Drmanac, Radoje T.
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Dickson, Mark
APPLICANT: Jones, Lee W.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
FILE REFERENCE: 774
CURRENT APPLICATION NUMBER: US/09/528,409
CURRENT FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/125,453
PRIOR FILING DATE: 1999-03-19

NUMBER OF SEQ ID NOS: 116231
SOFTWARE: HY-patent.pl Version 3.1
SEQ ID NO 59789
LENGTH: 422
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(422)
OTHER INFORMATION: n = A,T,C or G
US-09-528-409-59789

Query Match 85.2%; Score 23; DB 22; Length 422;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
|||||
Db 330 GGTGACTCACC GG GTGAACGGGG 352

RESULT 11
US-09-933-524-59789
Sequence 59789, Application US/09933524
GENERAL INFORMATION:
APPLICANT: Drmanac, Radoje T.
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Dickson, Mark
APPLICANT: Jones, Lee W.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
FILE REFERENCE: 774
CURRENT APPLICATION NUMBER: US/09/933,524
CURRENT FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 09/528,409
PRIOR FILING DATE: 2000-03-17
NUMBER OF SEQ ID NOS: 116231
SOFTWARE: HY-patent.pl Version 3.1
SEQ ID NO 59789
LENGTH: 422
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(422)
OTHER INFORMATION: n = A,T,C or G
US-09-933-524-59789

Query Match 85.2%; Score 23; DB 39; Length 422;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
|||||
Db 330 GGTGACTCACC GG GTGAACGGGG 352

RESULT 12
US-09-933-524A-59789
Sequence 59789, Application US/09933524A

GENERAL INFORMATION:
APPLICANT: Drmanac, Radoje T.
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Dickson, Mark
APPLICANT: Jones, Lee W.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
FILE REFERENCE: 774
CURRENT APPLICATION NUMBER: US/09/933,524A
CURRENT FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: 09/528,409

; PRIOR FILING DATE: 2000-05-17
 ; NUMBER OF SEQ ID NOS: 116231
 ; SOFTWARE: Hy-patent.pl Version 3.1
 ; SEQ ID NO 59789
 ; LENGTH: 422
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(422)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-933-524A-59789

Query Match 85.2%; Score 23; DB 39; Length 422;
 Best Local Similarity 100.0%; Pred. No. 8.9;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
 DB 330 GGTGACTCACC GG GTGAACGGGG 352

RESULT 13
 PCT-US02-22216-2
 ; Sequence 2, Application PC/TUS0222216
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of Virginia Patent Foundation
 ; APPLICANT: Chung, Ireland
 ; APPLICANT: Yeung, Fan
 ; TITLE OF INVENTION: Super Osteocalcin Promoter for the Treatment of Calcified Tumors
 ; FILE REFERENCE: 00704-02
 ; CURRENT APPLICATION NUMBER: PCT/US02/22216
 ; CURRENT FILING DATE: 2002-07-12
 ; PRIOR APPLICATION NUMBER: US 60/305,360
 ; PRIOR FILING DATE: 2001-07-13
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 565
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 PCT-US02-22216-2

Query Match 85.2%; Score 23; DB 2; Length 565;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
 DB 74 GGTGACTCACC GG GTGAACGGGG 96

RESULT 14
 PCT-US02-22216A-2
 ; Sequence 2, Application PC/TUS0222216A
 ; GENERAL INFORMATION:
 ; APPLICANT: The University of Virginia Patent Foundation
 ; APPLICANT: Chung, Ireland W. K.
 ; APPLICANT: Yeung, Fan
 ; TITLE OF INVENTION: Super Osteocalcin Promoter for the Treatment of Calcified Tumors
 ; FILE REFERENCE: 00704-02
 ; CURRENT APPLICATION NUMBER: PCT/US02/22216A
 ; CURRENT FILING DATE: 2002-07-12
 ; PRIOR APPLICATION NUMBER: US 60/305,360
 ; PRIOR FILING DATE: 2001-07-13
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 565
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

PCT-US02-22216A-2
 Query Match 85.2%; Score 23; DB 2; Length 565;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
 DB 74 GGTGACTCACC GG GTGAACGGGG 96

RESULT 15
 US-09-507-239-3
 ; Sequence 3, Application US/09507239
 ; GENERAL INFORMATION:
 ; APPLICANT: Keeping, Hugh S
 ; APPLICANT: Reichner, Jonathan S
 ; TITLE OF INVENTION: Treatment for Bone Disorders
 ; FILE REFERENCE: 21486-028
 ; CURRENT APPLICATION NUMBER: US/09/507,239
 ; CURRENT FILING DATE: 2001-02-18
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 597
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-507-239-3

Query Match 85.2%; Score 23; DB 22; Length 597;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
 DB 65 GGTGACTCACC GG GTGAACGGGG 87

Search completed: November 29, 2004, 16:15:10
 Job time : 3816 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 13:45:01; Search time 2981 Seconds
(without alignments)
330.048 Million cell updates/sec

Title: US-10-073-064-2

Perfect score: 27

Sequence: 1 GGTACTCACCAGCGGTGAACGGGCATT 27

Scoring table: IDENTITY_NUC

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsa1:
9: gb_gsa2:

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	85.2	436	7	R34738
2	19.6	72.6	495	1	AL904250
3	19.6	72.6	539	1	AL927557
4	19.6	72.6	776	9	CR131630
5	19.2	71.1	1150	2	BF026229
6	19	70.4	313	6	CD587090
7	19	70.4	352	6	CD587012
8	19	70.4	529	1	AI765212
9	19	70.4	569	8	BH474984
10	19	70.4	699	8	BH976919
11	19	70.4	725	8	BH517746
12	19	70.4	842	6	CA496206
13	18.6	68.9	716	7	CF879332
14	18.6	68.9	771	6	CB906760
15	18.6	68.9	790	5	BM368969
16	18.6	68.9	855	9	CNS004P9
17	18.6	68.9	976	7	CK865132
18	18.6	68.9	976	4	BM333940
19	18.6	68.9	1079	4	BM800594
20	18.6	68.9	1577	3	CR668008
21	18.2	67.4	406	7	CK554185
22	18.2	67.4	450	8	BZ096975
23	18.2	67.4	463	2	AW069362
24	18.2	67.4	540	4	BU689916

25	18.2	67.4	557	9	CE248534	CE248534	tigr-gsa-
26	18.2	67.4	622	7	CK535339	CK535339	rsmb0.00
27	18.2	67.4	675	9	CC711406	CC711406	OGWET5EH
28	18.2	67.4	802	9	CNS02ADH	AL188414	Tetradon
29	18.2	67.4	852	2	BE546517	BE546517	601076509
30	18.2	67.4	2567	3	BC037402	BC037402	Homo sap1
31	18	66.7	221	1	CE046711	CE046711	tigr-gsa-
32	18	66.7	261	1	AA249500	AA249500	18111-seq
33	18	66.7	321	1	AL374743	AL374743	MBB08H05
34	18	66.7	326	1	AL376677	AL376677	MBB08H06
35	18	66.7	335	9	CC645393	CC645393	OGWAM44TV
36	18	66.7	360	1	AV659678	AV659678	AV659678
37	18	66.7	418	1	AL376678	AL376678	MCB825D06
38	18	66.7	419	1	AL388075	AL388075	MCB46E02
39	18	66.7	427	1	AL388076	AL388076	MCB46E02
40	18	66.7	436	4	BI271584	BI271584	NE100G1F
41	18	66.7	461	1	AL376435	AL376435	MCB823H04
42	18	66.7	464	1	AL376436	AL376436	MCB823H04
43	18	66.7	468	5	BX472073	BX472073	DKF2686B
44	18	66.7	472	1	AL370632	AL370632	MCBA39A04
45	18	66.7	474	5	BQ147306	BQ147306	NF039C04F

ALIGNMENTS

RESULT 1
LOCUS R34738 436 bp mRNA linear EST 02-MAY-1995
DEFINITION YG61805.r1 Soares infant brain IN1B Homo sapiens cDNA clone
IMAGE:37522 5', mRNA sequence.

ACCESSION R34738.1 GI:791639
VERSION R34738
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Hillier, L., Clark, N., Dubuque, T., Eilston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The Wasnu-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

Insert Size: 2249
High quality sequence stops: 329 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2249 Std Error: 0.00
Seq primer: M18P1
High quality sequence stop: 329.
Location/Qualifiers

FEATURES

source

1. 436
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="Gene:410063"
/db_xref="taxon:9606"
/clone="IMAGE:37522"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain IN1B"
/note="Organ: whole brain; Vector: Lactid BA; Site 1: Not
I; Site 2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5

ACTGGAAGATCGCGCCGACAGATTTTCTTTTCTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lambda B4 vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 85.2%; Score 23; DB 7; Length 436;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGGCGTGAACGGCG 23
DB 304 GGTGACTCACCAGGCGTGAACGGCG 326

RESULT 2

AL904250/c AL904250 495 bp mRNA linear EST 02-JUL-2004
DEFINITION AL904250 PUR-Z1+22 Danio rerio cDNA clone 013-E07-1, mRNA sequence.

ACCESSION AL904250
VERSION AL904250.1 GI:23166809
KEYWORDS EST

SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 495)
Lo, J., Lee, S., Xu, M., Liu, F., Ruan, H., Eun, A., He, Y., Ma, W.,
Wang, W., Wen, Z., and Peng, U.
15000 unique zebrafish EST clusters and their future use in
microarray for profiling gene expression patterns during
embryogenesis

TITLE

JOURNAL Genome Res. 13 (3), 455-466 (2003)

MEDLINE 22505427
PubMed 12618376

COMMENT

Contact: Peng U
Lab of Functional Genomics
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore, 117609, Singapore
Email: pengj@imcb.a-star.edu.sg
Clone requests: info@openbiosystems.com
Open Biosystems,
6705 Odyssey Drive, Huntsville, AL 35806.

FEATURES

source

1..495
Location/Qualifiers
/organism="Danio rerio"
/mol_type="mRNA"
/strain="local wildtype"
/db_xref="taxon:7955"
/clone="013-E07-1"
/tissue_type="whole embryo or fish"
/dev_stage="mixed stages"
/clone_lib="PUR-Z1+22"

ORIGIN

Query Match 72.6%; Score 19.6; DB 1; Length 495;
Best Local Similarity 84.6%; Pred. No. 3e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGGCGTGAACGGCGCAT 26
DB 261 GGTGACCCACCGCGGTAAACGGGGGAT 236

RESULT 3

AL927557/c AL927557 539 bp mRNA linear EST 06-JUL-2004
LOCUS AL927557 PUR-Z1+22 Danio rerio cDNA clone 191-E01-1, mRNA sequence.

ACCESSION AL927557
VERSION AL927557.1 GI:23198469

ORIGIN

Query Match

EST.
Danio rerio (zebrafish)
Danio rerio

SOURCE

organism

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 539)
Lo, J., Lee, S., Xu, M., Liu, F., Ruan, H., Eun, A., He, Y., Ma, W.,
Wang, W., Wen, Z., and Peng, U.
15000 unique zebrafish EST clusters and their future use in
microarray for profiling gene expression patterns during
embryogenesis

AUTHORS

Lo, J., Lee, S., Xu, M., Liu, F., Ruan, H., Eun, A., He, Y., Ma, W.,
Wang, W., Wen, Z., and Peng, U.

TITLE

JOURNAL Genome Res. 13 (3), 455-466 (2003)

MEDLINE 22505427
PubMed 12618376

COMMENT

Contact: Peng U
Lab of Functional Genomics
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore, 117609, Singapore
Email: pengj@imcb.a-star.edu.sg
Clone requests: info@openbiosystems.com
Open Biosystems,
6705 Odyssey Drive, Huntsville, AL 35806.

FEATURES

source

1..539
Location/Qualifiers
/organism="Danio rerio"
/mol_type="mRNA"
/strain="local wildtype"
/db_xref="taxon:7955"
/clone="191-E01-1"
/tissue_type="whole embryo or fish"
/dev_stage="mixed stages"
/clone_lib="PUR-Z1+22"

ORIGIN

Query Match 72.6%; Score 19.6; DB 1; Length 539;
Best Local Similarity 84.6%; Pred. No. 3e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGGCGTGAACGGCGCAT 26
DB 392 GGTGACCCACCGCGGTAAACGGGGGAT 367

RESULT 4

CR131630/c CR131630 776 bp DNA linear GSS 06-JUL-2004
DEFINITION Forward strand read from insert in 3'HPT insertion targeting and
chromosome engineering clone MHP259023, genomic survey sequence.

ACCESSION CR131630
VERSION CR131630.1 GI:43879083
KEYWORDS GSS; genome survey sequence; MICE.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 776)
Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,
Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,
Rogers, J., and Bradley, A.

AUTHORS

Direct Submission
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICE

TITLE

JOURNAL CB10 1SA, UK. http://www.sanger.ac.uk/MICE

FEATURES

source

1..776
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP259023"
/clone_lib="MHP"

ORIGIN

Query Match 72.6%; Score 19.6; DB 9; Length 776;

Best Local Similarity 84.6%; Pred. No. 3.1e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GTGACTCAGCGGTGAACGGGCATT 27
DB 207 GTAACCTACCGGTGAACCGGCATT 182

RESULT 5

BF026229

LOCUS 601672804F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3955633 5',
DEFINITION mRNA sequence.

ACCESSION BF026229
KEYWORDS EST.
SOURCE GI:10733941

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1150)
NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at: image.lnl.gov
Plate: L1CM830 row: h column: 02
High quality sequence stop: 393.
Location/Qualifiers

FEATURES

source

1..1150
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3955633"
/tissue_type="melanotic melanoma"
/lab_host="NIH MGC 20"
/clone_lib="NIH MGC 20"
/note="Organ: skin; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 71.1%; Score 19.2; DB 2; Length 1150;
Best Local Similarity 87.5%; Pred. No. 4.8e+02;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GTGACTCAGCGGTGAACGGGCATT 25
DB 562 GAGACTCCCCGGGTGAACCGGCATT 585

RESULT 6

CD587090

LOCUS 313 bp mRNA linear EST 16-JUN-2003
DEFINITION RK039A1H1.1 T3 zebrafish Kidney Marrow cDNA library Danio rerio cDNA
clone RK039A1H1 5', mRNA sequence.

ACCESSION CD587090
VERSION CD587090.1 GI:31768442
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 313)

AUTHORS Song,H.D., Wu,X.Y., Sun,X.J., Zhou,Y., Liu,T.X., Zhang,G.W.,
Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y., Look,A.T.,
Zou,L.I. and Chen,Z.
TITLE Gene Expression Profiling in the zebrafish Kidney Marrow Tissue
JOURNAL Unpublished (2003)
CONTACT: Chen Z.

State Key Lab for Medical Genomics
Shanghai Institute of Hematology, Ruijin Hospital Affiliated to
Shanghai Second Medical University
197 Rui Jin Road II, Shanghai 200025, P. R. China
Tel: 86-21-64740430
Fax: 86-21-64743206
Email: zchen@smu.sh.cn
Seq primer: T3.

FEATURES

source

1..313
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="RK039A1H1"
/dev_stage="mature"
/note="Organ: kidney; Vector: pBS-CMV; Site 1: XhoI;
Site 2: EcoRI; Total RNA was extracted from the kidney
tissues of mature zebrafish. The poly (A)+ RNA fraction
was separated from total RNA by oligo (dT) cellulose
chromatography. Library was initially constructed in the
lambdaZAP Express vector (Stratagene) and in vivo excised
into pBS-CMV vector."

ORIGIN

Query Match 70.4%; Score 19; DB 6; Length 313;
Best Local Similarity 81.5%; Pred. No. 5.5e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGTGACTCAGCGGTGAACGGGCATT 27
DB 197 GTTACCGCCCCGTGAACCGGCATT 223

RESULT 7

CD587012

LOCUS 352 bp mRNA linear EST 16-JUN-2003
DEFINITION RK039A1H1.1 T3 zebrafish Kidney Marrow cDNA library Danio rerio cDNA
clone RK039A1H1 5', mRNA sequence.

ACCESSION CD587012
VERSION CD587012.1 GI:31768364
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 352)
Song,H.D., Wu,X.Y., Sun,X.J., Zhou,Y., Liu,T.X., Zhang,G.W.,
Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y., Look,A.T.,
Zou,L.I. and Chen,Z.
TITLE Gene Expression Profiling in the zebrafish Kidney Marrow Tissue
JOURNAL Unpublished (2003)
CONTACT: Chen Z.

State Key Lab for Medical Genomics
Shanghai Institute of Hematology, Ruijin Hospital Affiliated to
Shanghai Second Medical University
197 Rui Jin Road II, Shanghai 200025, P. R. China
Tel: 86-21-64740430
Fax: 86-21-64743206
Email: zchen@smu.sh.cn
Seq primer: T3.

FEATURES

source

1..352
Location/Qualifiers

/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="RK039A1H11"
/dev_stage="mature"
/clone_lib="zebrafish kidney marrow cDNA library"
/note="Organ: kidney; Vector: pBS-CMV; Site 1: XhoI;
Site 2: EcoRI; Total RNA was extracted from the kidney
tissues of mature zebrafish. The poly (A)+ RNA fraction
was separated from total RNA by oligo (dt) cellulose
chromatography. Library was initially constructed in the
lambdaZAP Express vector (Stratagene) and in vivo excised
into pBS-CMV vector."

ORIGIN

Query Match 70.4%; Score 19; DB 6; Length 352;
Best Local Similarity 81.5%; Pred. No. 5.5e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCAGGAGACGGGGCATT 27
Db 197 GTTACGCCCGCTGTGAACGGGCGCATT 223

RESULT 8 529 bp TRNA linear EST 28-JUN-1999
A1765212
LOCUS wi22f04.x1 NCI CGAP Kid12 Homo sapiens cDNA clone IMAGE:2398879 3
DEFINITION similar to SW:PAC_HUMAN Q00597 PANCONI ANEMIA GROUP C PROTEIN ;
mRNA sequence.

ACCESSION A1765212.1 GI:5231721

VERSION A1765212
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 529)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnl.gov/bdip/image/image.html
Seq primer: -40UP from Glbco
High quality sequence stop: 461.
Location/Qualifiers

FEATURES

SOURCE

1..529
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2398879"
/issue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid12"
/note="Organ: kidney; Vector: pTTT3-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kids was
prepared, and as circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."

ORIGIN

Query Match 70.4%; Score 19; DB 1; Length 529;
Best Local Similarity 81.5%; Pred. No. 5.7e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCAGGAGACGGGGCATT 27
Db 124 GGTGACTCTCCGTGTATCTGGGCATT 98

RESULT 9

BH474984 699 bp DNA linear GSS 13-DEC-2001

LOCUS BOGFA32TF BOGF Brassica oleracea genomic clone BOGFA32, genomic
survey sequence.

ACCESSION BH474984.1 GI:17683095

VERSION GSS.

KEYWORDS

SOURCE Brassica oleracea
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 699)
Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)

JOURNAL COMMENT Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: edcrom@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TP
Class: sheared ends.

FEATURES

source

1..699
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGFA32"
/clone_lib="BOGF"
/note="Vector: pROS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pROS1 using BstXI linkers"

ORIGIN

Query Match 70.4%; Score 19; DB 8; Length 699;
Best Local Similarity 81.5%; Pred. No. 5.7e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCAGGAGACGGGGCATT 27
Db 191 GGTGATTCGCCGAGAGAGGGGCTTT 217

RESULT 10

BH976919 699 bp DNA linear GSS 02-OCT-2002

LOCUS oth11f07.g1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.

ACCESSION BH976919.1 GI:23459922

VERSION GSS.

KEYWORDS

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 699)
Dehnanun, X., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

TITLE
JOURNAL
COMMENT
Nash, W., Rabinowicz, P.D. and Wilson, R.K.
Whole genome shotgun reads from Brassica oleracea
Unpublished (2002)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.wustl.edu
Plate: oehl row: f column: 07
Seq primer: -28Rport reverse

FEATURES
source
1. .699
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B. oleracea002"
/note="Vector: pOTW13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea T01000D3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN

Query Match 70.4%; Score 19; DB 8; Length 699;
Best Local Similarity 81.5%; Pred. No. 5.7e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCCTGTAACGGGGCATT 27
Db 426 GGTGATTCCTGGAGAACGGGGCTTT 452

RESULT 11
BHS17746/c 725 bp DNA linear GSS 13-DEC-2001
LOCUS
DEFINITION
BOGYE61TF BOGY Brassica oleracea genomic clone BOGYE61, genomic
survey sequence.
BHS17746
BHS17746.1 GI:17725936

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 725)
Town, C.D., Van Aken, S., Uterback, T., Koo, H. and Frazer, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOGYE61TR
Contact: Chris Town

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

FEATURES
source
1. .725
Location/Qualifiers

/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000D3"
/db_xref="taxon:3712"
/clone_lib="BOGYE61"
/clone_1lb="BOGY"
/note="Vector: pHOSt1, Site 1, BstXI; 2-3 kb sheared
genomic DNA inserted into pHOSt1 using BstXI linkers"

ORIGIN

Query Match 70.4%; Score 19; DB 8; Length 725;
Best Local Similarity 81.5%; Pred. No. 5.7e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCCTGTAACGGGGCATT 27
Db 226 GGTGATTCCTGGAGAACGGGGCTTT 200

RESULT 12
CA496206 842 bp mRNA linear EST 14-NOV-2002
LOCUS
DEFINITION
AGENCOURT 10811996 NCI CGAP ZK1d1 Danto rerio cDNA clone
IMAGE:6791831 5', mRNA sequence.
CA496206
CA496206.1 GI:24959287

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Danto rerio (zebrafish)
Danto rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danto.
1 (bases 1 to 842)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM414286 row: b column: 22
High quality sequence stop: 641.

FEATURES
source
1. .842
Location/Qualifiers

/organism="Danto rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6791831"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI CGAP_ZK1d1"
/note="Organ: Kidney; Vector: pCMV-SPORT6.1; Site 1:
EcoRV; Site 2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.8 kb. Constructed by J.
Wang (Research Genetics, Invitrogen Corp) from tissue
donated by L. Zon (Harvard University). Note: this is a
NCI CGAP Library."

ORIGIN

Query Match 70.4%; Score 19; DB 6; Length 842;
Best Local Similarity 81.5%; Pred. No. 5.8e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GGTGACTCACCCTGTAACGGGGCATT 27
Db 110 GGTGACCGCCCTGTGATGGGGCATT 84

RESULT 13
CF879332/c 716 bp mRNA linear EST 31-OCT-2003
LOCUS
DEFINITION
trich080xd08.b1 T.reesei mycelial culture, Version 6 October 2003
Hypocrea jecorina cDNA clone trich080xd08, mRNA sequence.
CF879332
ACCESSION
VERSION
KEYWORDS
EST.
Hypocrea jecorina (anamorph: Trichoderma reesei)

ORGANISM

Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

REFERENCE

1 (bases 1 to 716)

Dienert, S.E., Dunn-Coleman, N., Foreman, P., Houfek, T.D.,
Teunissen, P.U.M., van Solingen, P., Dankmeyer, L., Mitchell, T.K.,
Ward, M. and Dean, R.A.

TITLE

Characterization of the protein processing and secretion pathways
in a comprehensive set of expressed sequence tags from Trichoderma
reesei

JOURNAL

FEMS Microbiol. Lett. 230 (2), 275-282 (2004)

COMMENT

Contact: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024
Email: ralph.dean@ncsu.edu

FEATURES

source

Location/Qualifiers

1..716

/organism="Hypocrea jecorina"

/mol_type="mRNA"

/strain="Qm6a"

/db_xref="taxon:51453"

/clone="tric080xd08"

/dev_stage="mycelia"

/clone_lib="T. reesei mycelial culture, Version 6 October

2003"

/note="Vector: pREP3; Site 1: Not I/Sal I; Mycelial

culture grown from 24 hrs to 6 days with varying Carbon

and Nitrogen sources and concentrations."

ORIGIN

Query Match

Best Local Similarity 68.9%; Score 18.6; DB 7; Length 716;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db

1 GGTGACTCACCAGGCTGAACGGGGCA 25
43 GGTGACGACGCGGGTGCACGGGGAA 19

RESULT 14

CB906760

tric080xd08 T. reesei mycelial culture, Version 3 april Hypocrea

jecorina cDNA clone tric080xd08, mRNA sequence.

CB906760

EST.

CB906760.1 GI:30121418

EST.

Hypocrea jecorina (anamorph: Trichoderma reesei)

Hypocrea jecorina

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.

1 (bases 1 to 716)

Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Dienert, S.,

Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J.,

Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,

Olivera, H.A., Teunissen, P.U., Yao, J. and Ward, M.

Transcriptional regulation of biomass-degrading enzymes in the

filamentous fungus Trichoderma reesei

J. Biol. Chem. 278 (34), 31988-31997 (2003)

22803314

12788920

Contact: Pamela K. Foreman

Genencor Intl.

925 Page Mill Road, Palo Alto, CA 94304, USA

Tel: (650) 846-7635

Fax: (650) 621-7817

Email: pforeman@genencor.com

Seq primer: UP-F1 primer.

Location/Qualifiers

source

1..771

/organism="Hypocrea jecorina"

/mol_type="mRNA"

/strain="Qm6a"

/db_xref="taxon:51453"

/clone="tric080xd08"

/dev_stage="mycelia"

/clone_lib="T. reesei mycelial culture, Version 3 april"

/note="Vector: pREP3; Site 1: Not I/Sal I; Mycelial

culture grown from 24 hrs to 6 days with varying Carbon

and Nitrogen sources and concentrations."

ORIGIN

Query Match

Best Local Similarity 68.9%; Score 18.6; DB 6; Length 771;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db

1 GGTGACTCACCAGGCTGAACGGGGCA 25
98 GGTGACGACGCGGGTGCACGGGGAA 74

RESULT 15

CB368969

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On May 8, 2003 this sequence version replaced gi:30447801.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seque@genoscope.cns.fr, Web: www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

10183.X

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?cs=CS0BAE029ZF02_AE02703_2&c=10183.

FEATURES

source

Location/Qualifiers

1..790

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODL010YC17"

/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"

/cell_line="RAMOS CELL LINE"

/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT

25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match

Best Local Similarity 68.9%; Score 18.6; DB 5; Length 790;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db

1 GGTGACTCACCAGGCTGAACGGGGCA 25

Tue Nov 30 08:47:09 2004

Db 410 GGTGATTTCACCGGGGCGACGGGGCA 386

Search completed: November 29, 2004, 15:11:27
Job time : 2986 secs

us-10-073-064-2.rst

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2004, 13:45:01 ; Search time 1663 Seconds

(without alignments)
767.782 Million cell updates/sec

Title: US-10-073-064-2

Perfect score: 27

Sequence: 1 GGTGACTCACCAGGTGAACGGGGCATT 27

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4526729 seqs, 2364494745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

GenEmbl:*

- 1: gb_ba:*
- 2: gb_atg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pac:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_pro:*
- 11: gb_scs:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	100.0	27	6	AR009136 Sequence
2	27	100.0	27	6	AR202296 Sequence
3	23	85.2	57	6	AX100465 Sequence
4	23	85.2	57	6	AX100465 Sequence
5	23	85.2	597	6	EL1404 Regulatory
6	23	85.2	597	6	AX226068 Sequence
7	23	85.2	831	9	AY147065 Homo sapi
8	23	85.2	884	9	RATOSTRO
9	23	85.2	1384	12	AB105218 Mus sapien
10	23	85.2	1386	12	AB105219 Mus muscu
11	23	85.2	2650	12	AB004306 Artificial
12	23	85.2	2657	6	AX577978 Sequence
13	23	85.2	164168	9	AL135927 Human DNA
14	23	85.2	164179	9	AC007227 Homo sapi
15	20.2	74.8	95687	5	BX511245 Zebrafish
16	20.2	74.8	180506	2	BX005403 Danto rer
17	20.2	74.8	275351	2	AC117883 Ratcus no
18	19.8	73.3	267431	2	AC142799 Macaca mu
19	19.6	72.6	13203	6	CO614669 Sequence

C 20	19.6	72.6	154730	3	AC009184	AC009184 Drosophi
C 21	19.6	72.6	168062	2	AC012857	AC012857 Drosophi
C 22	19.6	72.6	172479	3	AC006590	AC006590 Drosophi
C 23	19.6	72.6	228644	10	AC102570	AC102570 Mus muscu
C 24	19.6	72.6	260027	3	AE003659	AE003659 Drosophi
C 25	19.6	72.6	263051	2	CR450811	CR450811 Danto rer
C 26	19.6	72.6	263051	2	CO597323	CO597323 Sequence
C 27	19.6	70.4	12531	6	CO596429	CO596429 Sequence
C 28	19.6	70.4	104334	5	AL068751	AL068751 Zebrafish
C 29	19.6	70.4	115940	9	AC119429	AC119429 Homo sapi
C 30	19.6	70.4	141855	5	BX122537	BX122537 Danto rer
C 31	19.6	70.4	152485	5	AL935211	AL935211 Zebrafish
C 32	19.6	70.4	153829	5	BX004864	BX004864 Zebrafish
C 33	19.6	70.4	154123	2	AC019562	AC019562 Drosophi
C 34	19.6	70.4	154143	2	BX502008	BX502008 Danto rer
C 35	19.6	70.4	158143	2	AL157384	AL157384 Human DNA
C 36	19.6	70.4	163815	9	AC138024	AC138024 Homo sapi
C 37	19.6	70.4	169686	2	BX57325	BX57325 Danto rer
C 38	19.6	70.4	169710	2	BX927124	BX927124 Danto rer
C 39	19.6	70.4	172741	5	AL954866	AL954866 Zebrafish
C 40	19.6	70.4	173298	2	AC006489	AC006489 Drosophi
C 41	19.6	70.4	176521	3	AC097725	AC097725 Drosophi
C 42	19.6	70.4	177426	2	CR456622	CR456622 Danto rer
C 43	19.6	70.4	181079	2	AC016667	AC016667 Homo sapi
C 44	19.6	70.4	184499	5	BX005201	BX005201 Zebrafish
C 45	19.6	70.4	184499	2	AC021209	AC021209 Homo sapi

ALIGNMENTS

RESULT 1	AR009136	Sequence 2 from patent US 5756264.	27 bp	DNA	linear	PAT 04-DEC-1998
LOCUS	AR009136					
DEFINITION	Sequence 2 from patent US 5756264.					
ACCESSION	AR009136					
VERSION	AR009136.1	GI:3967941				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	Unclassified.					
AUTHORS	1 (bases 1 to 27)					
TITLE	Schwartz,R.J., DeMayo,F.J. and O'Malley,B.W.					
JOURNAL	Expression vector systems and method of use					
FEATURES	Patent: US 5756264-A 2 26-MAY-1998;					
source	location/Qualifiers					
	1..27					
	/organism="unknown"					
	/mol_type="unassigned DNA"					

ORIGIN

Query Match	100.0%	Score 27;	DB 6;	Length 27;
Best Local Similarity	100.0%	Pred. No. 1.4;		
Matches	27;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
QY	1	GGTACTCACCAGGTGAACGGGGCATT	27	
Db	1	GGTACTCACCAGGTGAACGGGGCATT	27	
RESULT 2	AR202296	Sequence 2 from patent US 6361984.	27 bp	DNA
LOCUS	AR202296			
DEFINITION	Sequence 2 from patent US 6361984.			
ACCESSION	AR202296			
VERSION	AR202296.1	GI:20256835		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 27)			
TITLE	Cloeser,T., Ullrich,A. and Milauer,B.			
	MDK1 polypeptides			

JOURNAL Patent: US 6361984-A 2 26-MAR-2002;
 FEATURES Location/Qualifiers
 source 1..27
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 27; DB 6; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACC GG GTGAACGGGGCAT 27
 |||||
 Db 1 GGTGACTCACC GG GTGAACGGGGCAT 27

RESULT 3
 AX100465 57 bp DNA linear PAT 10-APR-2001
 LOCUS Sequence 6 from Patent WO0121649.
 AX100465
 ACCESSION
 VERSION AX100465.1 GI:13619490
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Adams,J.S.
 TITLE Nucleic acids encoding vitamin d response element binding proteins,
 products related thereto, and methods of using same
 JOURNAL Patent: WO 0121649-A 6 29-MAR-2001;
 CEDARS-SINAI MEDICAL CENTER (US)

FEATURES
 source 1..57
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 85.2%; Score 23; DB 6; Length 57;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACC GG GTGAACGGGG 23
 |||||
 Db 3 GGTGACTCACC GG GTGAACGGGG 25

RESULT 4
 AX100465/c 57 bp DNA linear PAT 10-APR-2001
 LOCUS Sequence 6 from Patent WO0121649.
 AX100465
 ACCESSION
 VERSION AX100465.1 GI:13619490
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Adams,J.S.
 TITLE Nucleic acids encoding vitamin d response element binding proteins,
 products related thereto, and methods of using same
 JOURNAL Patent: WO 0121649-A 6 29-MAR-2001;
 CEDARS-SINAI MEDICAL CENTER (US)

FEATURES
 source 1..57
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 85.2%; Score 23; DB 6; Length 57;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACC GG GTGAACGGGG 23
 |||||
 Db 56 GGTGACTCACC GG GTGAACGGGG 34

RESULT 5
 E13404 597 bp DNA linear PAT 27-APR-1998
 LOCUS Regulatory region of human osteocalcin gene.
 E13404
 ACCESSION
 VERSION E13404.1 GI:3252209
 KEYWORDS JP 1997182586-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Negishi,Y., Kajiyama,N. and Kawaguchi,R.
 TITLE 1 (bases 1 to 597)
 JOURNAL RECOMBINANT VECTOR, CELL TRANSFORMED THEREBY, AND MEASUREMENT OF
 BONE METABOLISM CONTROL ACTIVITY OF VITAMIN D DERIVATIVE USING THE
 Patent: JP 1997182586-A 1 15-JUL-1997;
 S R L:KK
 OS Homo sapiens (human)
 PN JP 1997182586-A/1
 PD 15-JUL-1997
 PF 29-DEC-1995 JP 1995352817
 PI NEGISHI YOICHI, KAJIYAMA NAOKI, KAWAGUCHI RYUJI PC
 C12N15/09,C12N5/10,C12P21/02,G01N33/82,(C12P21/02,C12R1:91); CC
 strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 FH Key Location/Qualifiers
 FH source 1..597
 FT /organism="Homo sapiens"
 FT /cell_type="osteoblast"
 FT /clone="OST-BETA-Gal"
 FT promoter 1..597
 FT /note="human osteocalcin gene promoter".

FEATURES
 source 1..597
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 85.2%; Score 23; DB 6; Length 597;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTGACTCACC GG GTGAACGGGG 23
 |||||
 Db 65 GGTGACTCACC GG GTGAACGGGG 87

RESULT 6
 AX226068 597 bp DNA linear PAT 10-SEP-2001
 LOCUS Sequence 3 from Patent WO0160983.
 AX226068
 ACCESSION
 VERSION AX226068.1 GI:15555393
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Keeping,H.S. and Reichner,J.S.

TITLE Treatment for bone disorders
JOURNAL Patent: WO 0160983-A 3 23-AUG-2001;
Rode Island Hospital, A Lifespan Partner (US)
FEATURES
source
1. .597
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 85.2%; Score 23; DB 6; Length 597;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
|||||
Db 65 GGTGACTCACC GG GTGAACGGGG 87
|||||

RESULT 7
LOCUS AY147065 831 bp DNA linear PRI 16-SEP-2002
DEFINITION Homo sapiens osteocalcin gene, promoter sequence.
ACCESSION AY147065
VERSION AY147065.1 GI:22947853
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 831)
AUTHORS Li, L.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2002) Institute of Orthopaedics, Xijing Hospital,
17# Changge West Road, Xi'an, Shaanxi 710032, P.R. China
FEATURES
source
1. .831
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>831
/gene="osteocalcin"
<1..>831
/gene="osteocalcin"
ORIGIN

Query Match 85.2%; Score 23; DB 9; Length 831;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
|||||
Db 329 GGTGACTCACC GG GTGAACGGGG 351
|||||

RESULT 8
LOCUS R405218 884 bp DNA linear PRI 28-FEB-2003
DEFINITION Homo sapiens osteocalcin gene, promoter region and partial cds.
ACCESSION M34013
VERSION M34013.1 GI:205865
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 884)
AUTHORS Morrison, N.A., Shine, J., Fragonas, J.C., Verkest, V., McMenamy, M.L.
and Eisman, J.A.
TITLE 1,25-dihydroxyvitamin D-3 responsive element and glucocorticoid
JOURNAL Science 246 (4934), 1158-1161 (1989)
MIDLINE 90069588

PUBMED 2588000
REFERENCE 2 (bases 1 to 884)
AUTHORS Morrison, N.A., Shine, J., Fragonas, J.C., Verkest, V., McMenamy, M.L.
and Eisman, J.A.
TITLE Direct Submission
JOURNAL Submitted (27-APR-1993) Garvan Institute of Medical Research, St.
Vincent's Hospital, Sydney, Australia.
FEATURES
source
1. .884
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1. .881
/note="contains promoter and 5' UTR"
misc_feature
<882..>884
/product="osteocalcin"
882..>884
/codon_start=1
/product="osteocalcin"
/protein_id="AAF05868.1"
/db_xref="GI:6223673"
/translation="M"
ORIGIN

Query Match 85.2%; Score 23; DB 9; Length 884;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACC GG GTGAACGGGG 23
|||||
Db 324 GGTGACTCACC GG GTGAACGGGG 346
|||||

RESULT 9
LOCUS AB105218 1384 bp DNA linear SYN 24-JAN-2004
DEFINITION Mus musculus transgenic DNA, hOsf1 transgene left flanking region,
integrated synthetic HpoC-Hosf-1 transgene.
ACCESSION AB105218
VERSION AB105218.1 GI:34740152
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
AUTHORS Tsujimura, A. and Hashimoto-Gotoh, T.
TITLE Transgene integration sites of hOsf1 expression mouse
JOURNAL Unpublished
2 (bases 1 to 1384)
AUTHORS Tsujimura, A. and Hashimoto-Gotoh, T.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-2003) Atsushi Tsujimura, Kyoto Pref. Univ. of Med.,
Dept. Biochem. & Mol. Genet. Res. Inst. for Neuro. Dis. & Geriatr.,
Kawaramachi-Hirokoji, Kajii-Cho 465, Kamigyo-Ku, Kyoto, Kyoto
602-8566 Japan (E-mail: atsujio@koto.kpu-m.ac.jp,
Tel:81-75-251-5850, Fax:81-75-251-5793)
FEATURES
source
1. .1384
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/transgenic
501..1384
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
<1..>500
/strand name="gamma satellite"
/note="hOsf1 transgene left flanking region"
/evidence="experimental"
501..>1384
/note="HpoC-Hosf-1 transgene"

ORIGIN /evidence=experimental

Query Match 85.2%; Score 23; DB 12; Length 1384;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGGTGAACGGGG 23
|||||
DB 840 GGTGACTCACCAGGTGAACGGGG 862

RESULT 10
LOCUS AB105219/C
DEFINITION Mus musculus transgenic DNA, integrated synthetic HpoC-Hosf-1
transgene, hosf1 transgene right flanking region.
ACCESSION AB105219
VERSION AB105219.1 GI:34740153
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 Tsujimura, A. and Hashimoto-Gotoh, T.
TITLE Transgene integration sites of Hosf1 expression mouse
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1386)
Tsujimura, A. and Hashimoto-Gotoh, T.
DIRECT SUBMISSION
JOURNAL Submitted (10-MAR-2003) Atsushi Tsujimura, Kyoto Pref. Univ. of Med.,
Dept. Biochem. & Mol. Genet. Res. Inst. for Neuro. Dis. Geriatr.;
Kawarachi-Hirokoji, Kajii-Cho 465, Kamigyō-Ku, Kyoto, Kyoto
602-8566, Japan (E-mail: atsujikoto.kpu-m.ac.jp,
Tel:81-75-251-5850, Fax:81-75-251-5799)

FEATURES
SOURCE Location/Qualifiers
1..1386
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/transgenic
1..886
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
complement(1..886)
/note="HpoC-Hosf-1 transgene"
complement(87..11386)
/standard_name="gamma satellite"
/note="hosf1 transgene right flanking region"
/evidence=experimental

ORIGIN

Query Match 85.2%; Score 23; DB 12; Length 1386;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGGTGAACGGGG 23
|||||
DB 545 GGTGACTCACCAGGTGAACGGGG 523

RESULT 11
LOCUS AB004306
DEFINITION Artificial construct HpoC-Hosf-1 transgene for osteoblast
stimulating factor-1, complete cds.
ACCESSION AB004306
VERSION AB004306.2 GI:18642488
KEYWORDS osteoblast stimulating factor-1.
SOURCE synthetic construct
ORGANISM synthetic construct

REFERENCE 1 artificial sequences.
AUTHORS Masuda, H., Tsujimura, A., Yoshioka, M., Arai, Y., Kuboki, Y., Mukai, T.,
Nakamura, T., Tsuji, H., Nakagawa, M., and Hashimoto-Gotoh, T.
TITLE Bone mass loss due to estrogen deficiency is compensated in
transgenic mice overexpressing human osteoblast stimulating
factor-1
JOURNAL Biochem. Biophys. Res. Commun. 238 (2), 528-533 (1997)
MEDLINE 97446018
PUBMED 9299545

REFERENCE 2 (bases 1 to 2650)
Hashimoto-Gotoh, T.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (27-MAY-1997) Tamotsu Hashimoto-Gotoh, Kyoto Pref. Univ.
of Medicine, RINDG, Dept. of Biochem. and Mol. Genet.,
Kawarachi-Hirokoji, Kyoto 602, Japan
(E-mail: tng@koto.kpu-m.ac.jp, Tel:075-251-5850, Fax:075-251-5799)
COMMENT On Feb 8, 2002 this sequence version replaced gi:2564035.
FEATURES
SOURCE Location/Qualifiers
1..2650
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="human osteocalcin promoter-driven human osf-1 gene
flanked by rabbit beta-globin 5'- and 3'- untranslated
regions; HpoC-Hosf-1 transgene used to generate transgenic
mice"
15..882
/note="HpoC-Hosf-1 transgene
derived from human osteocalcin promoter"
TATA_signal
CDS 823..828
1542..2048
/note="HpoC-Hosf-1 transgene
derived from human osf-1 gene"
/codon_start=1
/transl_table=1
/product="osteoblast stimulating factor-1"
/protein_id="BA2294.1"
/db_xref="GI:2564040"
/translation="MQAQYQOORRKAFLAFLAAYDTAEGAKKEKPEKVKK
SDGEMQWVCVPTSGDGTREGTGTGECOTMTORTKIPCNKKQPGACCKY
FOAMECDLNTALKRTNLSLRALHNACQTVTISRPGCLTPKPOAESKKKKKK
KKQKEMLD"

ORIGIN

Query Match 85.2%; Score 23; DB 12; Length 2650;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGACTCACCAGGTGAACGGGG 23
|||||
DB 343 GGTGACTCACCAGGTGAACGGGG 365

RESULT 12
LOCUS AX577978
DEFINITION Sequence 100 from Patent WO02081745.
ACCESSION AX577978
VERSION AX577978.1 GI:27647186
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Garcia, T., Roman Roman, S., Baron, R., Call, K., Theilhaber, J.,
Connolly, T., Jackson, A., Bushnell, S. E. and Rawadi, G.
TITLE Genes involved in osteogenesis, and methods of use
JOURNAL Patent: WO 02081745-A 100 17-OCT-2002;
Aventis Pharma S.A. (FR)
FEATURES
SOURCE Location/Qualifiers
1..2657

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 85.2%; Score 23; DB 6; Length 2657;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAGCTCACC GGGTGAACGGG 23
|||||
Db 343 GGAGCTCACC GGGTGAACGGG 365

RESULT 13

AL135927 164168 bp DNA linear PRI 30-SEP-2000
LOCUS Human DNA sequence from clone RPI1-54H19 on chromosome 1, complete
DEFINITION
SEQUENCE.

ACCESSION AL135927 GI:10443354
VERSION AL135927.14
KEYWORDS HTG.

SOURCE Homo sapiens (human)

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 164168)
Howden, P.

REFERENCE
TITLE Direct Submission
JOURNAL Submitted (27-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk

COMMENT

On Oct 1, 2000 this sequence version replaced gi:10185404.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RPI1-54H19 is from the library RPI1-11.1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://pacpac.med.buffalo.edu/
VECTOR: pACE3.6
This sequence is the entire insert of clone RPI1-54H19 The true
right end of clone RPI1-17216 is at 50458 in this sequence.

FEATURES

source

1..164168
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RPI1-54H19"
/clone_lib="RPI1-11.1"

ORIGIN

Query Match 85.2%; Score 23; DB 9; Length 164168;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAGCTCACC GGGTGAACGGG 23
|||||
Db 129362 GGAGCTCACC GGGTGAACGGG 129384

RESULT 14
AC007227/c 164179 bp DNA linear PRI 20-NOV-1999
LOCUS Homo sapiens chromosome 1 clone RPI1-11_54H19, complete sequence.
DEFINITION
AC007227
AC007227.3 GI:6456148

VERSION
KEYWORDS HTG.

SOURCE Homo sapiens (human)

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 164179)
Jones, D., Mundt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
Jones, M., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
Bryant, J., Tesmer, J., Meincke, L., Longmire, J., White, S., Tatam, O.,
Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
Han, C. and Deaven, L.

Sequencing of Human Chromosome 1
Unpublished

REFERENCE
AUTHORS 2 (bases 1 to 164179)
Ricke, D.O.

Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
Unpublished
3 (bases 1 to 164179)

REFERENCE
AUTHORS Bruce, D., Mundt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
Jones, M., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
Bryant, J., Tesmer, J., Meincke, L., Longmire, J., White, S., Tatam, O.,
Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
McMurry, K., Han, C. and Deaven, L.

REFERENCE
AUTHORS Direct Submission
JOURNAL Submitted (06-APR-1999) Center for Human Genome Studies, DOE Joint
Genome Institute, Los Alamos National Laboratory, MS M888, Los
Alamos, NM 87545, USA
4 (bases 1 to 164179)

REFERENCE
AUTHORS Bruce, D., Mundt, M., Doggett, N., Munk, C., Saunders, E., Robinson, D.,
Jones, M., Buckingham, J., Chasteen, L., Thompson, S., Goodwin, L.,
Bryant, J., Tesmer, J., Meincke, L., Longmire, J., White, S., Tatam, O.,
Campbell, C., Fawcett, J., Maltbie, M., Bussod, M., Sutherland, R.,
McMurry, K., Han, C. and Deaven, L.

REFERENCE
AUTHORS Direct Submission
JOURNAL Submitted (20-NOV-1999) Center for Human Genome Studies, DOE Joint
Genome Institute, Los Alamos National Laboratory, MS M888, Los
Alamos, NM 87545, USA
On Nov 20, 1999 this sequence version replaced gi:5103801.

COMMENT
LOCATION/Qualifiers

FEATURES

source

1..164179
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RPI1-11_54H19"
44..354
/rpt_family="Alu"
175..474
/rpt_family="Alu"
366..665
/rpt_family="Alu"
921..1071
/rpt_family="Alu"
misc_feature
5041803..1889,3009..3122,4079..4159,8103..8192,
8296..8385,8783..9025,9799..9939,10100..10699,
12907..13098,15854..16030,17315..17536,23403..23561,
23976..24068,25002..25079,25486..25623,25808..25891)
/note="X1A protein (Homo sapiens) BAA63041.1"
2627..2766
/rpt_family="Alu"

```

repeat_region 3617. .3907
/rpt_family="Alu"
repeat_region 4568. .4914
/rpt_family="Alu"
repeat_region 5455. .5755
/rpt_family="Alu"
repeat_region 6422. .7051
/rpt_family="Alu"
repeat_region 7069. .7507
/rpt_family="Alu"
repeat_region 7541. .7717
/rpt_family="MIR"
repeat_region 9550. .9679
/rpt_family="Alu"
repeat_region 11237. .11737
/rpt_family="Alu"
repeat_region 11237. .11737
/rpt_family="Alu"
repeat_region 11209
/rpt_family="MIR2"
repeat_region 112187. .12512
/rpt_family="Alu"
repeat_region 11383. .14286
/rpt_family="Alu"
repeat_region 14310. .14594
/rpt_family="Alu"
repeat_region 14590. .14933
/rpt_family="Alu"
misc_feature 15109. .15184
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 1650. .16947
/rpt_family="Alu"
repeat_region 17672. .17965
/rpt_family="Alu"
repeat_region 19255. .19569
/rpt_family="Alu"
repeat_region 20077. .20164
/rpt_family="Alu"
repeat_region 20560. .20876
/rpt_family="Alu"
repeat_region 21341
/rpt_family="Alu"
repeat_region 21364. .21673
/rpt_family="Alu"
repeat_region 21737. .22047
/rpt_family="Alu"
misc_feature 22972. .23038, 23401. .23562, 23974. .24053
/rpt_family="Alu"
/note="NCI CGAP Kid1 Homo sapiens cDNA (AM000913)"
join(23974. .24059, 25001. .25079, 25481. .25537)
/note="BSTR865 Cerebellum II Homosapiens cDNA (AA325805)"
25847. .25930
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 27587. .27873
/rpt_family="Alu"
repeat_region 30267. .30394
/note="GRAIL 2 excellent exon, frame 2"
join(30610. .60686, 31242. .31321, 31559. .31721, 32071. .32187)
/note="Stratagene schizo Brain s11 Homo sapiens (AA63560)"
complement(33070. .34744)
/note="Human gene for bone gla protein (BGP). Accession number X04143"
repeat_region 33332. .33453, 33649. .33720
/note="Human osteocalcin precursor (gamma-carboxyglutamic acid). Accession number F02818"
complement(34258. .35135)
/note="Human osteocalcin gene, 5' end, and promoter region (M34013)"
repeat_region 35719. .36177
/rpt_family="Alu"
misc_feature 36901. .37900, 39601. .40600, 41345. .41500,
complement(join(36901. .37900, 44800, 45881. .46000, 49501. .49782, 52201. .53200))
/note="Homo sapiens polyamine modulated factor-1 (AF141309)"
repeat_region 38224. .38381

```

```

repeat_region 38438. .38781
/rpt_family="L1"
repeat_region 38735. .39038
/rpt_family="Alu"
repeat_region 39252. .39367
/rpt_family="Alu"
repeat_region 39395. .39520
/rpt_family="L1"
repeat_region 40725. .41022
/rpt_family="Alu"
repeat_region 41056. .41317
/rpt_family="Alu"
repeat_region 41393. .41667
/rpt_family="MIR2"
repeat_region 41673. .41977
/rpt_family="Alu"
repeat_region 42395. .42527
/rpt_family="Alu"
repeat_region 42562. .42704
/rpt_family="Alu"
misc_feature 42789. .42889
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 43007. .43300
/rpt_family="Alu"
repeat_region 43395. .43782
/rpt_family="Alu"
misc_feature 44092. .44197
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 44770. .45050
/rpt_family="Alu"
repeat_region 45070. .45248
/rpt_family="Alu"
repeat_region 45264. .45564
/rpt_family="Alu"
repeat_region 45567. .45701
/rpt_family="Alu"
repeat_region 46910
/rpt_family="Alu"
repeat_region 47258. .47576
/rpt_family="Alu"
misc_feature 48074. .48187
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 48208. .48545
/rpt_family="Alu"
repeat_region 48512. .48678
/rpt_family="Alu"
repeat_region 48665. .48996
/rpt_family="Alu"
repeat_region 49023. .49273
/rpt_family="Alu"
repeat_region 49717. .49867
/rpt_family="Alu"
repeat_region 50215. .50644
/rpt_family="MIR"
repeat_region 51016. .51174
/rpt_family="L1"
repeat_region 51301. .51581
/rpt_family="Alu"
repeat_region 51693. .51987
/rpt_family="Alu"
repeat_region 52848. .53012
complement(52848. .53012)

```

```

Query Match 85.2% Score 23; DB 9; Length 164179;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTAAGTACCGGCGGACGGGG 23
Db 34807 GGTAAGTACCGGCGGACGGGG 34785
RESULT 15

```


BX511245

LOCUS

BX511245 95687 bp DNA linear VRT 21-JUL-2004

DEFINITION

zebrafish DNA sequence from clone CH211-125016 in linkage group 17,

complete sequence.

ACCESSION

BX511245 GI:50470853

VERSION

KEYWORDS

SOURCE

ORGANISM

Danio rerio (zebrafish)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Emi.: EMBL; Sw.: SWISSPROT; Tr.: TREMBL; Wp.: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TR repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhixiong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, Washu). For further information see

http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml

CH211-125016 is from a CHORI-211 BAC library

VECTOR: pPARBAC2.1.

Location/Qualifiers

1..95687

/organism="Danio rerio"

/mol_type="genomic DNA"

/db_xref="taxon:7855"

/clone="CH211-125016"

/clone_11b="CHORI-211"

ORIGIN

Query Match 74.8% Score 20.2; DB 5; Length 95687;

Best Local Similarity 88.0%; Pred No. 2.7e-02;

Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 TGACTCACCAGGTGAACGGGGCATT 27

DB 10650 TGACACACCGGTGTGAATGGGGCATT 10674

